LECTURES ON THE SCIENTIFIC BASIS OF MEDICINE
1953–54

British Postgraduate Medical Federation
University of London

LECTURES ON THE SCIENTIFIC BASIS OF MEDICINE

Volume III

1953–54

UNIVERSITY OF LONDON
THE ATHLONE PRESS
1955

Published by
THE ATHLONE PRESS
UNIVERSITY OF LONDON
at 2 Gower Street, London, W.C.1
Distributed by Constable & Co. Ltd.
12 Orange Street, London, W.C.2

U.S.A.
John de Graff, Inc.
64 West 23rd Street,
New York, 10

Printed in Great Britain by
WESTERN PRINTING SERVICES LTD
BRISTOL

PREFACE

THERE is no need to preface this, the third annual volume of *Lectures on the Scientific Basis of Medicine*, with an account of the origin and purpose of these Lectures: they have been stated fully in the prefaces to the previous volumes. As in previous years this volume fairly reflects the content of the particular series of lectures from which its material is drawn. The lectures here published should, like the earlier ones, assist graduates to appreciate the advances which are being made in fields other than those in which they are personally engaged and which may affect the development of the subjects in which they are chiefly interested.

The present volume reflects, especially, the growing knowledge of the physical and chemical structure of biological tissues, of how these are built up and maintained, of their reactions to injury and their interdependence as a factor contributing to the health of the individual. The use of radioisotopes and the method of chromatography have enabled advances to be made in many directions. The actions of hormones and enzymes and the importance of the connective tissues are again emphasized. The elucidation of the processes of infection and immunity, to which the discovery of penicillin gave new impetus, and the ever widening field of chemotherapy, are dealt with in a number of the lectures The chemical transmission of nerve impulses and the actions of the anticholinesterases are again represented in this volume. Two subjects, experimental psychopathology and dermatology, appear for the first time and this is perhaps indicative of the relatively early stage that has been reached in the knowledge of the scientific basis of these branches of medicine, an appreciation of the opportunities for fruitful work in these fields may, it is hoped, stimulate the search for new knowledge

The first lecture, by Dr. H. E. Sigerist, on 'Science and History' offers justification, if such is needed, for the publication of this and other volumes which endeavour to assist communication over a broad field of knowledge.

A fourth volume based on the series of lectures for 1954–55 is already in preparation, and will, it is hoped, be published early in 1956.

FRANCIS R. FRASER
Director, British Postgraduate Medical Federation

7 January 1955

NOTE

The lectures printed in this volume were delivered on the following dates.

I	26 November 1953	XI.	5 November 1953
II.	15 October 1953	XII	10 November 1953
III	26 January 1954	XIII.	4 March 1954
IV.	19 January 1954	XIV.	4 February 1954
V.	2 March 1954	XV	11 February 1954
VI.	14 January 1954	XVI.	9 February 1954
VII	1 December 1953	XVII.	2 February 1954
VIII.	20 October 1953	XVIII.	16 February 1954
IX.	22 October 1953	XIX.	28 January 1954
X	19 November 1953	XX.	3 December 1953
XXI.	13 October 1953		

CONTENTS

I

Science and History

HENRY E. SIGERIST

So far the lectures in this series have been dealing with various aspects of modern science as the foundation of modern medicine. May I be allowed to look at science from a somewhat different angle? Science and history seem to have very little in common. Few scientists are interested in history and modern science is so absorbing that it leaves little time for studies in other fields. A scientist, who formerly was well known, once said to me that he was not interested in studying history because he was making history. He is dead now and unfortunately very little of his work has stood the test of time. The historian on the other hand has a very scanty knowledge of science, not much more than what he learned in secondary school. And yet the two, science and history, have much more in common than appears on the surface. The sharp division between science and the humanities is artificial, a late development, the result of specialization. It did not exist in the Renaissance. Girolamo Fracastoro, the 400th anniversary of whose death was commemorated widely in 1953, was a physician, to be sure, who made important contributions to our knowledge of contagion and contagious diseases, but he was a humanist first and foremost who wrote poetry and whose range of interests was extraordinarily wide. His contemporary, Paracelsus, was a physician also, but he wrote theological and philosophical works. And even in the eighteenth century a [illegible]

period, entered the literary scene with a volume of poems which

exerted a very strong influence on German literature and in some ways even anticipated Goethe. Besides writing enormous tomes on botany, anatomy, physiology, besides editing many works of students and colleagues and writing innumerable letters on scientific subjects, Haller also wrote three novels and several theological books. The great French philosophers of the period of enlightenment—men like Diderot, d'Alembert, Rousseau, Voltaire—were philosophers, poets, historians and were all keenly interested in science.

The great development of the various sciences in the nineteenth and particularly in the twentieth centuries was responsible for the split between science and the humanities and we are sometimes inclined to identify the humanities with culture and science with technology, which is quite wrong. The work of scientists like Darwin or Haeckel had greatly influenced the philosophic outlook of men of my generation. At a conference of the World Health Organization at Nancy in 1952 the general opinion was that medical students should be cultured, that they should have a broad humanistic outlook, and one member of our group made the very pertinent remark that science was also culture and that the training in the sciences that the medical students received was contributing to the development of such a broad humanistic outlook in them to a very great extent.

We all know that science has a future, and probably a very great future. The progress achieved in the few decades since my graduation from medical school in 1917 has been stupendous. We are, however, all too inclined to forget that science also has a past and that today's progress is the outcome and result of a very long development. The historian, on the other hand, is only too ready to forget that science is one of the major factors in the moulding of history. I should like to discuss the subject of science and history from two angles. First and foremost I would like to point out the tremendous significance of the time factor in the development of science, and secondly I will say a few words about science as an element in the making of history.

A poet who writes an elegy, a form of poetry very popular in antiquity, or a sonnet, a form of poetry brought to great perfection in the Renaissance, or a poem in free verse, may have

written a very good poem, but it need not necessarily be better than an elegy of Archilochus or a sonnet of Petrarch. In other words there is no progress in poetry as there is in science. Within the framework of the society of which he is a member the poet expresses the fears and hopes, the joys and anxieties that he and many of his contemporaries feel and does it in a given style which is the style of his period. He does not have to build on the accumulated experience of the centuries, but the scientist does. An observation may be correct, but it is wasted if the time is not ripe and the foundations are missing. Let me give you a very simple example from the medical literature of ancient Greece.

The Hippocratic physicians observed that in certain cases of bronchitis you heard a distinctive murmur inside the chest which sounded like boiling vinegar. In the case of dry pleurisy they heard something that sounded as if a new leather strap was bent to and fro. In other cases of pleurisy they shook the patient—this method was later to be named Succussio Hippocratis—and heard a distinct sound. In other words the Greeks in the fifth century B.C. had found the principle of auscultation. Why did they not develop it? Why did the world have to wait for Laënnec before auscultation became a generally accepted method of examination? For the simple reason that the Greeks did not think in terms of pathological anatomy. Since the middle of the eighteenth century we have known that many symptoms of disease are the functional expression of anatomical changes in the organs, and therefore methods were developed so that these changes could be perceived with our sense organs in the living patient. In 1761 the Viennese physician Auenbrugger invented the method of percussion, and it was not by accident that this invention was made by the son of an innkeeper who used to knock at his kegs in order to hear whether they were full or empty. Nor was it an accident that he was a good musician, as was also Laénnec, because only people with a good ear were able to discover slight differences in sounds. Auenbrugger's [illegible] hought he [illegible] which we [illegible] Nostalgia, however, is an excellent example showing how medical views

are sometimes determined by outside factors and are time-bound.

An Alsatian medical student, Johannes Hofer, wrote a doctoral dissertation in 1611 at the University of Basel under the title *De Nostalgia oder Heimwehe* in which he described nostalgia as a disease entity of its own. The dissertation met with immediate success, was reprinted and translated. J. J. Scheuchzer took up the subject and attributed responsibility for the disease to the air. It was for very good reasons that Swiss physicians were particularly interested in this disease. At that time many Swiss soldiers were in foreign service and it happened occasionally that one of them became so homesick that he deserted. Such a highly dishonourable action was resented by the whole nation. If homesickness, however, were a physical ailment caused by changes of atmospheric pressure when a mountaineer went to live in the lowlands, then there was an excuse for desertion, disease being always considered as an accident which excused people from many of the obligations under which people in good health stood.[1]

Another example of a strange time-conditioned disease is that of tarantism. It occurred in Southern Italy, particularly in Apulia, in the Middle Ages and the Renaissance and we have descriptions of it even from the seventeenth and eighteenth centuries. The disease was attributed to the sting of a spider, the Tarantula. People were attacked by it at the height of the summer heat, in July and August. They suddenly jumped up, feeling an acute pain like the sting of a bee Some saw the spider, others did not, but they knew it must be the Tarantula. They ran out of the house into the street, to the market place, dancing in great excitement. Soon they were joined by others who like them, had just been bitten or by people who had been stung in previous years, for the disease remained in the body and was reactivated every year by the heat of summer. People were known to have relapsed every summer for thirty years All ages were affected, children as well as old people, although most of

[1] For the history of nostalgia see Fritz Ernst, *Vom Heimweh*, Zürich 1949. An English translation of Hofer's dissertation was published in *Bull Inst Hist Med* 2, 376 (1934)

them were men and women in the prime of life. More women than men were attacked by the disease. Its victims were mostly peasant people, but ladies and gentlemen and even worthy monks and nuns were not spared. People danced wildly in the queerest attire, dressed in strange costumes with necklaces, in dresses of bright colour, red, green and yellow, but they could not endure the sight of black. Some would tear their clothes and show their nakedness, losing all sense of modesty. They waved red cloths in their hands, wore wreaths of vine leaves and waved boughs of vine. Some called for swords and acted like fencers, others for whips and flagellated each other, women called for mirrors and howled making indecent motions. Others again liked to be tossed in the air while still others rolled in the dirt like swine. They all drank wine plentifully and acted and talked like drunken people. We owe these details to a monograph written by an outstanding Italian physician, Giorgio Baglivi, who studied the disease on the spot.[1]

There was only one cure for the disease and this was music and dancing. Bands of musicians roamed the country at the height of summer playing the tarantella, repeating the tunes an endless number of times until people broke down perspiring profusely, whereupon they were cured, at least for that year, but the following year the sound of the tarantella would reactivate the poison that was believed to be in their system. We know the music that was played, as the learned Jesuit father Athanasius Kircher collected the tunes and published them in his work *Magnes sive de arte magnetica*, in 1641. Baglivi being a good iatromechanist explained the action of the poison mechanically but as early as the eighteenth century it was found that the sting of the Tarantula was perfectly harmless or at least not more harmful than the sting of a bee or a wasp. The explanation must be sought in a totally different direction. In 1621 a physician, Epiphanius Ferdinandus, put his finger on the right spot when he said that some people considered the disease *melancholiae seu amentiae quaedam species*, some kind of melancholy or insanity and

are sometimes determined by outside factors and are time-bound.

An Alsatian medical student, Johannes Hofer, wrote a doctoral dissertation in 1611 at the University of Basel under the title *De Nostalgia oder Heimwehe* in which he described nostalgia as a disease entity of its own. The dissertation met with immediate success, was reprinted and translated. J. J. Scheuchzer took up the subject and attributed responsibility for the disease to the air. It was for very good reasons that Swiss physicians were particularly interested in this disease. At that time many Swiss soldiers were in foreign service and it happened occasionally that one of them became so homesick that he deserted. Such a highly dishonourable action was resented by the whole nation. If homesickness, however, were a physical ailment caused by changes of atmospheric pressure when a mountaineer went to live in the lowlands, then there was an excuse for desertion, disease being always considered as an accident which excused people from many of the obligations under which people in good health stood.[1]

Another example of a strange time-conditioned disease is that of tarantism. It occurred in Southern Italy, particularly in Apulia, in the Middle Ages and the Renaissance and we have descriptions of it even from the seventeenth and eighteenth centuries. The disease was attributed to the sting of a spider, the Tarantula. People were attacked by it at the height of the summer heat, in July and August. They suddenly jumped up, feeling an acute pain like the sting of a bee. Some saw the spider, others did not, but they knew it must be the Tarantula. They ran out of the house into the street, to the market place, dancing in great excitement. Soon they were joined by others who like them, had just been bitten or by people who had been stung in previous years, for the disease remained in the body and was reactivated every year by the heat of summer. People were known to have relapsed every summer for thirty years. All ages were affected, children as well as old people, although most of

[1] For the history of nostalgia see Fritz Ernst, *Vom Heimweh*, Zürich 1949 An English translation of Hofer's dissertation was published in *Bull. Inst. Hist Med.* 2, 376 (1934).

been reliable enough. So they devised a very simple method, stirred a fine clay in water and applied this suspension rapidly to the back of the chest. The point where the clay dried first obviously was the hottest spot and there the incision was made.

Aristarchus in the third century B.C. taught at Alexandria that the sun is at rest, that the earth rotates about its own axis and that the earth like other planets circles around the sun. Today we know that he was right, but the time was not ripe for his discovery. Nobody believed him, he was charged with impiety and the geocentric system of Ptolemy dominated astronomical theories until the time of Copernicus.

Another example which I think illustrates very graphically the importance of the time element in science is to be found in the work of the German anatomist and pathologist Jakob Henle. In 1840 he published a remarkable book on pathological investigations the first part of which discusses miasmas and contagium, and miasmatic-contagious diseases. He demonstrates that miasma is a substance which enters the human body from the outside world and causes disease. The prototype of such a disease is malaria which is always acquired from outside and not by contact. A contagium on the other hand is produced in the body and is transmitted from one individual to another. The prototype of such a disease is syphilis, which is transmitted by contact. Most other infectious diseases, however, are miasmatic-contagious. An individual acquires a miasma from outside, develops a contagium which then is passed on from one to the other. Such a disease is the plague. One gets it from outside through a miasma which then multiplies in the body and is passed on to large numbers of individuals. If, however, miasma and contagium can produce the same disease they must be identical, and must be not only organic but live material which develops in the body like a parasite. Henle's treatise was a masterpiece of logic and was correct in all its basic assumptions, yet there was no response to it. Why? Again due to the time element. In 1840 German medicine was just liberating itself from the lofty speculations of *Naturphilosophie* and people wanted to see things and not merely believe in them. Twenty years later,

this neurosis becomes understandable when one remembers that it occurred in a region where orgiastic cults, the cults of Dionysus, Cybele, Demeter and others were celebrated, cults which had a decidely erotic character. Christianity came late to Apulia and came to a primitive and conservative population in which ancient beliefs and customs were deeply rooted. The pagan rites, so popular with the people, were now considered sinful. But they survived nevertheless, probably in secret, and one day they came to the light again as symptoms of a disease which not only excused, but legitimized them. It is one of many pagan institutions that survived in the Christian world after having changed meaning.

The observation of the Hippocratic physicians of murmurs within the chest in certain diseased conditions showed that a correct observation cannot be developed if the time is not yet ripe for it. If we go back to the Hippocratic writings we may find, however, that observation and correct reasoning could give good results even if the theory was wrong. Thus the Greeks knew that pneumonia did not always end in crisis but that sometimes empyema developed. According to the theory pneumonia was a phlegmatic disease, the phlegm became pus which accumulated in the pleural cavity and had to be driven out through the natural healing power of the body. The pus would then either break through the wall of the chest or into the bronchi. The Greeks knew that such a process took a very long time, so long in fact that the patient had a good chance of dying before the pus had broken through. The physician's task therefore was to help nature in its healing tendency, that is, to create an artificial opening for the pus to break through. The problem, however, was to find the pus, and to know where the incision should be made. To us this is very easy because we have the method of percussion, we have X-rays and we can make an exploratory puncture. The Greeks did not have any of these methods and they had to rely on reasoning. Empyema was an inflammatory disease. Hence heat was developed, and the point of greatest heat must be the point where the pus had collected and where the incision was to be made. But it was not easy to find that point. Putting the hand on the chest would not have

Renaissance painted his figures with clear outlines, well-defined surfaces, his composition was harmonious and well balanced. The Baroque artist, on the other hand, saw the world in motion, the outlines of his figures were erased, played upon by light and shadow. He emphasized the diagonal in his compositions, an open window, a distant landscape may catch your eye, and this gave his creation depth and wider perspective. His art was dynamic, not static. The Renaissance artist was interested in what exists, the Baroque artist in what happens. There was a definite change from a static to a dynamic outlook, one which we find not only in art, but in music beginning with Caccini, in physics with Galileo and in the medical sciences with Harvey. He was an anatomist, as we all know, but what fascinated him was motion and in his hands anatomy became *anatomia animata*. Harvey wrote another book which characteristically enough had embryology for a subject. Embryology also is dynamic anatomy, one that changes from one moment to another.

Every science requires a foundation upon which to build. There could be no scientific physiology without anatomy, and today physiology has to draw heavily from the discoveries made in physics and chemistry. The example of Harvey shows that a scientific movement can only develop in a certain atmosphere. It follows the general trends of the period, and it is perfectly obvious that the social and economic conditions prevailing at that time must have a strong influence on the development of science, acting either as a stimulus and incentive or as a retarding factor. Ancient Greece was a seafaring nation and as soon as their ships went beyond the Aegean sea some astronomical knowledge was required which in turn called for mathematics. Throughout history navigation has been a strong stimulus to science. The larger the ships became the further they went and the more scientific knowledge was required.

Slave economy, on the other hand, was a handicap to science. As long as slave labour was easily available and cheap there was no demand for labour-saving machines. The Greeks knew the principle of the steam engine but never applied it. There was no need for vastly increased industrial production in a society

Pasteur and Robert Koch were able to show under the microscope what the miasmas and contagiums actually were and thus satisfy the inquisitive spirit of a rationalist age.

In 1928 the tercentenary of the publication of William Harvey's book describing the circulation of the blood was commemorated all over the world. I was at the University of Leipzig at the time and was asked to give a formal address on the occasion. I did not want to repeat what had been so aptly said many times before, that Harvey was a brilliant investigator—there had been many others before and after him—that he made the experimental method one of the chief methods of research in biology, that he very successfully applied quantitative considerations. Experiments had been performed before Harvey, and one of his contemporaries Santorio Santorio tried to solve another biological problem quantitatively, the problem of metabolism. To that end he spent part of his life on scales weighing carefully what entered and what left his body, but the problem was one that could not be solved before the revolution in chemistry brought about by Lavoisier. It puzzled me to know why the circulation of the blood was discovered in the first half of the seventeenth century, why not before, why not later? There were many brilliant anatomists in the sixteenth century, and even before that time the lesser circulation had been seen incidentally by some and described, although its full significance was not recognized. If we wish to understand correctly a new scientific development, we must study it within the framework of the general civilization of the period, studying that civilization in all its aspects, economic, social, literary, artistic, etc. When we do this in the case of Harvey we soon find that at the end of the sixteenth century and in the early seventeenth century a basic change took place in man's outlook on the world. The relation between the individual and the world changed and he looked at it with different eyes. A new art developed with Michelangelo—an art which matured in the seventeenth century and which we call the art of the Baroque as

hopes of the eighteenth and nineteenth centuries when it was assumed that science would be the means of liberating the people once and for all from the bonds of disease, hunger and poverty.

The time element also noticeably affects the speed with which a discovery is accepted. Auenbrugger's percussion was a most valuable invention as it gave the physician an extremely useful diagnostic method, but hardly any notice was taken of it at the time his book was published and almost half a century passed before the method was seriously discussed and then gradually accepted. The authority of Corvisart, physician-in-ordinary to Napoleon, was needed to draw general attention to percussion. In 1808 he published a new edition of Auenbrugger's book with a French translation and a voluminous commentary, and from then on physicians began to practise percussion, in France first, then in England, and finally also in Auenbrugger's home land Austria, and in Germany. The explanation for this delayed recognition is easy to find. The year 1761 not only saw the publication of Auenbrugger's book, but also of Morgagni's great work *De sedibus et causis morborum per anatomen indagatis*, a book which was to become the foundation of pathological anatomy. Anatomical thinking was not yet general enough to warrant the acceptance of percussion.

In the seventeenth century Santorio constructed a thermometer to measure the temperature of patients suffering from fever. It was a glass ball that the patient kept in his mouth, connected with a tube on which the temperature could be read. But at that time nobody thought of measuring fever and for centuries physicians believed that putting the hand on the patient's forehead was enough to determine a sick man's fever, although Boerhaave in the early eighteenth century and de Haen some time later had used the thermometer in their clinics at Leiden and Vienna. The measuring of temperature as a matter of routine became general only in the second half of the nineteenth century. On the other hand, the discovery of the X-ray spread like wildfire. I will give you a few dates for France only but I am sure they were very similar in other countries. On the 28th December 1895 Wilhelm Roentgen announced his discovery. A few weeks later, on the 10th February 1896, Charles

which consisted for the most part of slaves, small farmers and craftsmen. Handicraft industry was perfectly sufficient to satisfy the needs of a small upper class. The very ingenious machines described by Hero of Alexandria remained to a large extent on paper. Conditions changed when the Roman Empire was pacified and slaves became rarer and more expensive, and even more so when slavery was abolished altogether in Christian Europe. There was a great shortage of labour in the Middle Ages and as a result great efforts were made to make better use of animal power, water power, and wind power.

In the Renaissance the demand for metals increased considerably. An increased volume of trade required more gold as a medium of exchange and voyages of discovery were launched in search of gold. It was found in Mexico, in Peru, and the early expeditions to North America were also undertaken in the search for gold. The new firearms demanded more copper and lead, and since the shallow deposits of minerals were exhausted new machinery was required and new health hazards created. It is not by accident that the first monographs on occupational diseases and particularly on miners' diseases were written at that time.

Men of my generation have experienced two world wars. Both brought endless destruction and suffering to the world, but we cannot deny that they were also a strong stimulus to science. We remember what aviation was before World War I and how tremendously it developed during and after the war. The second World War brought radar, the use of atomic energy, and speed in the application of the sulpha drugs, penicillin and DDT which would have been quite impossible without the pressure of war. War means destruction. But it is no fault of the scientists if their discoveries are used for destructive purposes. It is the fault of the people who have not yet learned to create the social organization that the new science requires. The *New Yorker*, that greatest of all American magazines, printed a highly significant cartoon a few years ago. A young man was telling his parents that he wanted to become a scientist, whereupon one of the parents, terrified, said, 'Isn't there enough trouble in the world already?' How far have we moved from the

poison is evil, but poison may also be a remedy. It is the dosage that determines the effect. The third sphere is *Ens naturale*. Men may be contemporaries but no two are identical and we know well enough that no two individuals have the same fingerprint or the same handwriting. Every individual is born with a nature of his own and to a large extent carries his destiny within himself. The fourth sphere is *Ens spirituale*. Man has body and mind like all other animals, they are one inasmuch as they determine one another mutually. But man is an animal of a special kind. He is conscious of himself and of his past, he not only feels pain but is able to reflect about the phenomenon of pain and to establish abstract concepts. Thus the spirit in Paracelsus' terminology gives man a special position in nature and from the spirit causes of disease may also arise.

Now let us go back to the first entity. What is *Ens astrale*? The stars move according to eternal laws, and so does man's life. The constellation characterizes a given moment and every individual has his moment, his historical time, which affects his life in health and disease. This is a very fine and correct thought. Thirty years ago I had an ordinary pneumococcus pneumonia which developed into an empyema and I was ill for many months. Today the same pneumonia would have been cured in a few days and the empyema avoided. People suffering from pernicious anaemia, diabetes, meningitis, erysipelas, puerperal fever and many other diseases have good chances for survival today, while they were lost only yesterday. On the other hand, men of my generation had good chances of being killed in two wars The historical moment in other words affects not only scientific developments, but also man's health and illness.

Let me make just a few remarks about science as a factor in the moulding of history. I shall be brief because the facts are more generally known, and I shall limit myself to the impact of science on society and to a few recent developments in the western world. As long as agriculture was primitive and industry consisted of craftsmen operating on a small scale, little science was needed and applied in the process of production. From the Renaissance onwards, however, scientific thought and

Henry presented a report on Roentgen's discovery to the French Academy of Science. On the 17th April the first French X-ray machines were shown. Before that, on the 1st April a clinical demonstration was made before the Société de Chirurgie and on the 6th August before the Congrès Français de Médecine.[1] Why such a difference in the acceptance of thermometry and of the X-ray? Again the reason is not difficult to find. Little was known about the physiology of fever before the nineteenth century and it took the physicians some time to think about it in quantitative terms. The use of X-rays for purposes of diagnosis, however, was on the straight line of medical development and was the crowning method which had been preceded by percussion, auscultation, the invention of the ophthalmoscope and of the laryngoscope, by apparatuses that introduced electric bulbs and mirrors into every cavity of the body in an attempt to see anatomical changes in the living organisms. The X-ray which permitted one to look through the body was eagerly accepted in a very short time. A factor which must be kept in mind also is that Roentgen never took a patent on his discovery, but even if he had the method would have been adopted without delay.

The time factor plays a part in another sense and, to illustrate this, I would like to go back to Paracelsus and to one of his most inspiring books, the *Volumen Paramirum*, a book on which he worked for many years and which he completed around 1530.[2] The book is not easy to read as it is written in the symbolic terms so dear to the author, but it is undoubtedly the ripest fruit of Paracelsus' thought. It is an attempt to outline a medical [illegible]

of nature, he lives in a physical environment from which he derives matter and energy. Food comes from nature but so does poison, there are normal but also abnormal stimulations. Everything that comes from nature is both good and evil, food is good,

[1] H. Péquignot, 'La médecine et le monde moderne', *Les Temps Modernes*, 9, 773 (1953).

[2] An English translation by K. F. Leidecker was published as Supplement No 11 to *Bull Hist Med*, Baltimore, 1949

parties. The polarization of society to the right and to the left is a process in the midst of which we still find ourselves today. The industrialization of the West also created a new political outlook. Sources of raw material were needed to feed the industries and foreign markets to absorb their products, thus leading every western country to endeavour to build a colonial empire.

In the field of biology Darwin's theory of evolution had a profound influence upon philosophic and religious thought, so much so that its teaching is still forbidden in parts of America. As a result of science, the general material standard of living was raised, at least in some western countries. Life has become safer, more comfortable and we have an infinity of enjoyable gadgets. Science is worshipped and there is no better recommendation for a toothpaste or breakfast cereal than to advertise it as being scientific. A man like Einstein is universally respected although very few people understand what his contribution to science has been. But he is highly esteemed as the embodiment and symbol of science.

Science, however, also knows of many frustrations. Industries apply scientific principles to their production, to be sure, but by no means freely. We all know of the endless number of patents that are purchased by firms in order to be suppressed, in order to forestall competition or because a firm wants to use up its old machinery. Governments do not act scientifically. They consult with scientists, obviously, hundreds or even thousands of scientists are in the government service as experts, and in the United States the Department of Agriculture is one of the foremost scientific research institutions of the country. But politics are the result of compromise, conflicts of vested interests are unavoidable. Foreign policy is still less scientific. If we could exploit the resources of the world scientifically, if production, distribution and consumption could be organized along scientific lines, the standard of living would be raised considerably. But we all know how difficult this problem is Socialism is an attempt to organize the life of a nation in a scientific manner and we shall have to see how successful it will be.

When, thirty years ago, some of us talked of the social implications of science nobody listened. In the current text-books of

scientific discoveries began to exert an ever increasing influence. The voyages of discovery in the sixteenth century made a deep impression on western society, new continents were found with plants, animals, and races of mankind unknown to the Greeks. The human body was explored by Vesalius and his fellow anatomists, and the universe by Copernicus and his followers. This discovery of the world broadened man's outlook considerably, and it was continued in the eighteenth century by Newton, Galileo, Kepler and many others. At the same time Harvey, Descartes, Borelli showed that the human body was some kind of a mechanical system and that mechanical laws applied to it also. The microscope revealed a world of infinitely small living beings. One discovery followed another. In the eighteenth century it was electricity. Chemistry had its great revolution and the steam engine was invented. The philosophers deeply interested in science became its chief popularizers and propagandists. The French Revolution enthroned reason, and reason is the essence of science.

In the course of the nineteenth century in Europe first, in America somewhat later, science became a determining factor in history. New industries developed, and industrial output both increased considerably in quantity and improved in quality. The populations of all industrial countries increased because new means of transportation made it possible to bring food from distant regions and the population could grow beyond the capacity of the home soil. At the end of the last and in the beginning of our own century science created new industries directly, such as the electrical, the chemical, the food and cosmetics industries, and a second industrial revolution took place which affected the western world as strongly as the first.

The industrialization of the West created a new social and economic order. The majority of people became wage earners or salaried employees who depended for their living on the labour market, over which they had no control. The working class organized itself in militant trade unions and political

II

Biological Synthesis

G. POPJÁK

Introduction

ONE of the most significant advances in biochemistry during the past twenty years has been the increasing realization of the important role small molecules play in the metabolism of the living world—both in the plant and animal kingdoms. Plants represent perhaps the extreme example of simplicity of taste for food, being able to build up their cell constituents, their proteins, fats, carbohydrates, their pigments—and the substances which have become the vitamins for the animal world—all from CO_2, ammonia, or nitrates, mineral salts, water and sunlight. While we ourselves are more fastidious in our choice of diet—we prefer steak to ammonium sulphate and we would rather have butter on our bread and sugar in our coffee than inhale CO_2 even on a gloriously sunny day—nevertheless when our cells and tissues are laying down their constituents they will more frequently resort to a synthetic process from small molecules in preference to accepting unchanged compounds from the diet or even substances absorbed from the blood. The process involves, of course, the necessity of first degrading large molecules and then reassembling the fragments into a specific pattern Such a mechanism has the effect of preserving the characteristic composition of the tissues of a particular species, and further, in the course of the stepwise reactions of degradation and resynthesis integration with the energy metabolism of the cells takes place, the whole integrated process being the maintenance of life with the performance of specific

history the word science was as a rule not even mentioned. The situation has changed today and we have many excellent textbooks of all grades which fully recognize the great part played by science in the historical development of the western world. The scientist, on the other hand, is beginning to be aware of his social responsibilities. He is not only an expert but has an important part to play as an expert citizen.

The history of science can teach us a great deal. In our schools science is too frequently taught in a dogmatic way and not presented in its cultural setting. A body of generally accepted knowledge, simplified and carefully digested, is transmitted to students who accept it as a matter of course. The graduate teaching of science trains specialists, highly efficient specialists, who, however, are frequently uneducated outside their speciality. The academic world surrendered so readily to dictatorship in many countries because it consisted of specialists who knew nothing outside the narrow bounds of their special field. If we wish to produce a citizen able to think in terms of science and a scientist prepared to participate in social action we must improve our methods of teaching. One way, and in my opinion it holds great promise, is to approach the sciences not only technically but also historically, philosophically and sociologically.

chemical bonds with the aid of highly efficient catalysts in the form of enzymes and various pigments.

The living world, as we know it, can be divided into two large groups: (1) into autotroph and (2) into heterotroph organisms. The autotroph organisms are characterized by their ability to build up their cell constituents entirely from inorganic matter and to assimilate carbon dioxide. Plants, green algae, green, red and purple bacteria—all autotrophs—use light quanta as the source of energy for the assimilation of CO_2. Another group of autotroph bacteria cannot make use of light energy; instead, they derive the energy for the same purpose from the oxidation of inorganic substances like hydrogen sulphide, ammonia, molecular hydrogen, or ferrous iron. Heterotroph organisms, which include the pathogenic bacteria and all animal species, on the other hand must be supplied with organic substances, made ultimately by the autotrophs, in order to live. Some heterotrophs, like bacteria and yeasts, can live, multiply, and build up all their constituents when grown on very simple media provided a suitable organic carbon source is added, e.g. lactate, acetate, ethanol, glycerol. In the hetero-

On the basis of the source of the energy which drives the primary synthetic reactions we speak (i) of photosynthesis and (ii) of chemosynthesis, as in the two classes of autotroph organisms, and (iii) of enzymic synthesis as it occurs in all heterotroph organisms. It must not be thought, however, that enzymic synthesis is a special prerogative of heterotrophs; autotroph organisms abound in enzymes and enzymic reactions too. A germinating seed contains several enzymes and is entirely heterotroph and depends for its existence on substances stored in the seed until after the formation of chlorophyll in the cotyledons, when its photosynthetic life can begin.

PHOTOSYNTHESIS

It is hardly necessary to emphasize that photosynthesis, whereby carbon dioxide is being assimilated into organic substances is

functions, such as secretion, transfer, muscular contraction, etc. The continuous breakdown and regeneration of most of our body constituents is now an established fact and is mediated through the reactions of small molecules such as acetate, pyruvate, oxaloacetate, α-ketoglutarate, amino acids and other substances. The processes by which living cells effect the building up of large molecules, proteins, fats and carbohydrates, from small ones is the subject of my lecture and is called biological synthesis or, for short, biosynthesis.

Relation of Synthesis to Energy-yielding Processes

As I have indicated, biosynthesis is closely integrated with energy-yielding processes and it is, therefore, impossible to discuss one without the other. Biosynthetic processes, and the synthetic reactions carried out in the test tube by the organic chemist are endergonic processes, i.e. they can take place only if the necessary energy for the reactions is available. Seemingly, there is a tremendous difference between the reactions of synthetic organic chemistry as practised by man and biosynthetic reactions as practised by the cells, in that the intermediates in the two kinds of synthesis aiming at the same product are usually widely different. Further, syntheses in the cell proceed generally at much milder conditions of temperature and hydrogen-ion concentration, and much more efficiently than the reactions of synthetic organic chemistry. No organic chemist can as yet claim to be able to convert, for example, some grams of CO_2 into carbohydrate in the course of a few hours in a nearly quantitative yield, as a few green tobacco leaves will do (cf. Porter and Martin, 1952); neither can he convert acetic acid into fatty acids with an efficiency of 25–40 per cent in four or five hours, as a few grams of liver slices can do when incubated under suitable conditions. Yet fundamentally there is a feature common to man-made synthetic chemistry and the synthetic chemistry of the cell, in that in both instances energy must be provided. The organic chemist employs heat, acids, and alkalis and what might in popular terms be called 'strong' reagents, to achieve his purpose. Where the cell scores over man is in the harnessing of energy derived from various sources into special

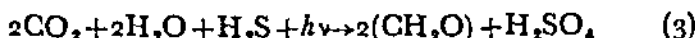
The evolution of oxygen as a 'waste' product is unique to plant photosynthesis and it was thought, therefore, for some time that this type of reaction differs in some fundamental way from the photosynthesis observed in some bacteria. However, a detailed study of the photosynthetic bacteria leads finally to a generalized concept of photosynthesis which puts all photosynthetic reactions on to a common basis.

From the investigation of photosynthesis in green bacteria —an anaerobic species living in media containing hydrogen sulphide—van Niel (see van Niel, 1941) concluded that the photosynthetic reactions in these bacteria can be represented by equation (2):



$$CO_2+2H_2S+h\nu\rightarrow(CH_2O)+H_2O+2S \qquad (2)$$

Or, in the purple sulphur-bacteria (Thiohordacea), which live in the same type of environment as the green bacteria, the photosynthetic reduction of CO_2 can be expressed by equation (3):



$$2CO_2+2H_2O+H_2S+h\nu\rightarrow2(CH_2O)+H_2SO_4 \qquad (3)$$

Another example of bacterial photosynthetic assimilation of CO_2 is that occurring with Rhodovibrio (Gaffron, 1935) which uses molecular hydrogen and is shown in equation (4):

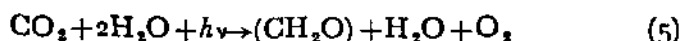


$$CO_2+2H_2+h\nu\rightarrow(CH_2O)+H_2O \qquad (4)$$

In all these reactions CO_2 is being reduced to carbohydrate and another [illegible]

substances which can serve as hydrogen donors in photosynthesis. The Thiohordacea can use for the same purpose also short-chain fatty acids, and the red non-sulphur bacteria (Athiohordacea) use exclusively organic substances of a large variety as oxidizable substrates in photosynthesis.

If the photosynthetic reaction (1) of plants be rewritten in the form (5):



$$CO_2+2H_2O+h\nu\rightarrow(CH_2O)+H_2O+O_2 \qquad (5)$$

there is a striking resemblance, for example, to equation (2)

the most fundamental of all life processes, as the entire heterotroph population of the world—which includes ourselves—depends on it for existence. It has been estimated that $1{\cdot}6 \times 10^{10}$ tons of carbon are fixed annually by plants in photosynthesis (cf. Gaffron, 1946), thus enormous amounts of energy are being harnessed in a readily utilizable form.

While, in this lecture, I would like to discuss mainly enzymic synthesis, which is of the more immediate import from the medical point of view, it will be necessary as an introduction to outline some of the main features of photosynthesis as well.

The over-all reaction of photosynthesis in green plants and algae is given by formula (1):



$$CO_2 + H_2O + h\nu \rightarrow (CH_2O) + O_2 \qquad (1)$$

where h is the Planck-constant and ν the frequency of absorbed light, their product representing the energy of light quanta. (CH_2O) is the general formula for a carbohydrate unit; in the most common of carbohydrates, in the hexoses, there are six such units. As is evident from formula (1) we depend on the photosynthetic life of plants not only because we draw from it our organic foodstuffs but also because it provides us with oxygen for respiration From the above photosynthetic reaction we can read off another characteristic, common to a large number of biosynthetic processes, viz. that it is a reductive process. The movement and transport of H_2 molecules among the cell constituents is the most fundamental of all energy transformations of life.

In photosynthesis the light energy is used not for the actual fixation of carbon dioxide but for the provision of activated hydrogen for the reductive process of the reactions. The carbon dioxide is fixed in the cells as *the carboxyl group of an already* existing organic 'acceptor' substance. This is an enzymic reaction and does not require light. The effect of light is to cause, through the activation of chlorophyll, a photolysis of water; then, through the intervention of enzymes the hydrogen of water is transferred to acceptor and donor substances which finally carry out the reduction of the carboxylated product of CO_2-fixation

temperature. It is now firmly established (Lipmann, 1941) that the utilizable energy obtained from the intermediary metabolism of foodstuffs is converted into a special chemical energy in the form of the 'energy-rich' pyrophosphate bond of adenosinetriphosphate (ATP). Investigations of recent years are gradually unfolding the mechanisms by which the potential energy of ATP is being used in biosynthetic reaction.

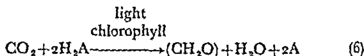
The concept of the 'energy-rich' phosphate bond has been one of the most fruitful ideas in biochemical research during the last decade. It originated with the discovery of creatine-phosphate by Eggleton and Eggleton (1927) and by Fiske and Subarrow (1929) and with the observation of Meyerhof and Suranyi (1927) that the enzymic hydrolysis of the phosphate bond in creatine-phosphate yielded the unexpectedly high energy of about 12,000 calories per mol of substance hydrolysed. The hydrolysis of a pyrophosphate bond of adenosine-triphosphate yields likewise about 12,000 calories. This is in contrast to the mere 2,000–4,000 calories released from the hydrolysis of an ordinary ester phosphate linkage. Hence the term 'energy-rich' phosphate bond for the former and 'energy-poor' phosphate bond for the latter type of linkage. It should be pointed out here that this biochemical concept of an 'energy-rich' bond does not coincide with the bond energies of the physico-chemist, viz. with the energy required to break a chemical bond, because our 'energy-rich' bonds can be easily split, i.e. in the physico-chemical sense these are weak bonds. The 'energy-poor' bonds of an ester phosphate linkage on the other hand can be broken with much greater difficulty, in other words these are strong bonds. The essential quality of an 'energy-rich' bond may be easily grasped if it is looked upon as carrying a high potential energy.

There are five types of 'energy-rich' bonds (*Figure 1*) which, according to a convention adopted after the suggestion of Lipmann (1941) are designated by the symbol $\sim$:

1. The guanidine-phosphate bond as in creatine and arginine phosphate;
2. The carboxyl-phosphate bond as in 1 : 3-diphosphoglyceric acid;

expressing the photosynthetic assimilation of CO_2 by green bacteria in the presence of hydrogen sulphide. Kluyver and Donker (quoted by Franck and Gaffron, 1941) expressed the view as long ago as 1926 that the water in plant photosynthesis should be looked upon as an oxidizable substrate, and the evolved oxygen as the oxidation product of water.

On the basis of these ideas van Niel (1941) formulated a general equation (6) for photosynthesis:



$$CO_2+2H_2A \xrightarrow[\text{chlorophyll}]{\text{light}} (CH_2O)+H_2O+2A \qquad (6)$$

which states that for the photosynthetic assimilation of CO_2 in the form of (CH_2O) a hydrogen donor (H_2A) is required which becomes oxidized to water and another molecular species (A), through the action of light on chlorophyll or 'bacterio-chlorophyll'.

The unique feature of plant photosynthesis, namely the evolution of oxygen, became now a matter of course, because only the photosynthetic reaction that used water as the oxidizable hydrogen donor could produce molecular oxygen. That the O_2 evolved comes truly from water was shown by the fact that isolated chloroplasts, which do not assimilate CO_2, when exposed to light in an aqueous medium containing oxidizing agents to act as hydrogen acceptors, produce molecular oxygen (Hill and Scarisbrick, 1940).

Enzymic Synthesis. Source of Energy for Synthesis in Animal Cells

Before I proceed to discussing enzymic synthesis, it will be appropriate to enquire first briefly into some problems of energy transformations in animal cells.

In animal cells, which cannot capture the energy of sunlight, energy derived from the metabolic breakdown of foodstuffs is utilized in biosynthetic reactions. However, energy liberated from the oxidation of foodstuffs in the form of heat is lost as such and cannot be made use of; the animal cell cannot work as a heat engine because its reactions occur virtually at an even

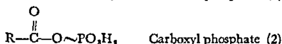
the mechanism of the formation of the pyrophosphate bond in adenosinetriphosphate. This subject has been reviewed admirably by Professor Krebs (1953). I can only indicate briefly the stages in the intermediary metabolism of foodstuffs at which the formation of such bonds takes place.

During the initial breakdown of complex carbohydrates, fats and proteins in the intestine (into hexoses, glycerol and fatty acids, and into amino acids), glycosidic- ester-, and peptide bonds are hydrolysed. The hydrolysis of these bonds yields, however, less than 1 per cent of the total free energy available from these three foodstuffs; this energy is lost entirely as heat. After absorption from the intestine, the primary breakdown products of foodstuffs enter a second stage of degradation in the course of intermediary metabolism within tissue cells and yield three substances: (1) acetate (from hexoses, glycerol, fatty acids and several amino acids); (2) α-ketogluterate (from amino acids); and (3) oxaloacetate (from amino acids). These three relatively simple substances enter the third and final stage of oxidative breakdown in the citric acid cycle.

Energy utilizable for synthetic reactions becomes available from both the second and third stages of foodstuff degradation (cf. Krebs, 1953). Two energy-rich phosphate bonds are generated in the course of anaerobic glycolysis: (i) in the conversion of 3-glyceraldehyde phosphate into 1 3-diphosphoglyceric acid, and (ii) in the conversion of 2-phosphoglyceric acid itno phospho-*enol*-pyruvic acid. The energy-rich phosphate bonds are then transferred to adenosinediphosphate to form ATP. Pyrophosphate bond of ATP is also generated during the oxidative decarboxylation of α-ketoacids (cf. Krebs, 1953). The most important sources of the energy-rich phosphate bonds of ATP are, however, the dehydrogenation steps of the citric acid cycle shown in Figure 2. During each revolution of the cycle four molecules of H_2 are removed at various stages. The energy from this oxidative cycle becomes available, not at the time of oxidation (dehydrogenation) of the members of the cycle, but when the hydrogen removed reacts—ultimately through the cytochrome system—with molecular oxygen to form water. The free energy available from this reaction is about 52,000

3. The enol-phosphate bond as in phospho-*enol*-pyruvic acid;
4. The pyrophosphate bond as in adenosinetriphosphate; and
5. The acyl-mercaptide bond as, for example, in acetyl-coenzyme A.





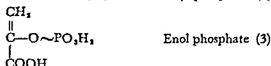






FIG. 1. Types of energy-rich bonds.

With the exception of the enol-phosphate all these bonds are essentially acid anhydrides; and you may easily convince yourselves of the 'energy-richness' of this type of bond by observing the evolution of much heat on adding to acetic anhydride a little water, which hydrolyses the anhydride bond. The panultimate of all these 'energy-rich' bonds is the pyrophosphate bond in ATP, as the others represent either intermediary stages in energy transformation during anaerobic glycolysis, as (2) and (3), or can be formed at the expense of the pyrophosphate bond of ATP, as in (1) and (5). Creatine-phosphate in vertebrates and arginine-phosphate in invertebrates represent, of course, a storage form of potential chemical energy in organs, such as muscle, nerve-tissue and electrical organs of fishes, which may be called upon suddenly to do intensive work.

It is beyond the scope of this lecture to give a detailed account of the energy transformations in the body or to discuss

man and their colleagues, and to Lynen, Reichert and Rueff (1951) in Germany. In particular, the latter workers have shown that the reactive grouping in coenzyme A (CoA) is the -SH radical; acetyl-CoA being an acyl-mercaptide containing an energy-rich bond mentioned earlier. Kaplan and Lipmann (1948) showed that acetyl-CoA was formed when either a liver or yeast enzyme was incubated with acetate, CoA and ATP. No acetyl-CoA is formed in the absence of ATP. The over-all reaction can be represented by the equation (7):



$$\text{acetate}+\text{HS-CoA}+\text{ATP}=\text{acetyl}\sim\text{S-CoA}+\text{AMP}+\text{PP} \quad (7)$$

according to which the energy-rich acyl-mercaptide bond is formed at the expense of one pyrophosphate bond of ATP, the latter giving rise to adenosine monophosphate (AMP=Ad-P) and inorganic pyrophosphate (PP) (Jones, Lipmann, Hilz and [illegible] more than one way. For example the enzyme could catalyse first the phosphorylation of HS-CoA with ATP (8):



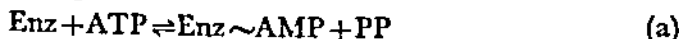
$$\text{Ad-P}\sim\text{PP}+\text{HS-CoA}=\text{PP}\sim\text{S-CoA}+\text{Ad-P} \quad (8)$$

and then catalyse the exchange of the pyrophosphate grouping with acetate (9):

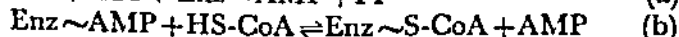


$$\text{PP}\sim\text{S}-\text{CoA}+\text{acetate}=\text{acetyl}\sim\text{S-CoA}+\text{PP} \quad (9)$$

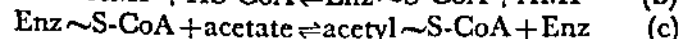
Indeed, Lipmann inclined to this view until very recently (cf. Lipmann, 1952). In a still later publication, however, Jones, Lipmann, Hilz and Lynen (1953) adduced experimental evidence to show that the transfer of the energy-rich bond from ATP to acetyl∼S-CoA involves the formation of enzyme-complexes containing the energy-rich bond and not the phosphorylation of CoA. Their results showed that the formation of acetyl∼S-CoA proceeds in all probability according to the following steps:



$$\text{Enz}+\text{ATP}\rightleftharpoons\text{Enz}\sim\text{AMP}+\text{PP} \quad (a)$$



$$\text{Enz}\sim\text{AMP}+\text{HS-CoA}\rightleftharpoons\text{Enz}\sim\text{S-CoA}+\text{AMP} \quad (b)$$



$$\text{Enz}\sim\text{S-CoA}+\text{acetate}\rightleftharpoons\text{acetyl}\sim\text{S-CoA}+\text{Enz} \quad (c)$$

The combination of the three steps (a), (b) and (c) gives the over-all reaction (7).

calories (Ball, 1944; Burton and Wilson, 1953) and can give rise to three pyrophosphate bonds of ATP (Lehninger, 1949; 1951; Lehninger and Smith, 1949; Slater, 1950).

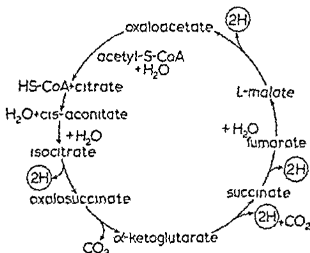


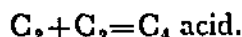
FIG. 2. Diagrammatic representation of the citric acid cycle.

ROLE OF ENERGY-RICH BONDS IN BIOSYNTHESIS

At present there are only a few biosynthetic reactions in which the role of ATP is known with any degree of precision. One of these is the formation of the so-called 'active' acetate with co-enzyme A (CoA). The important part acetate plays in biosynthetic reactions as the primary building unit for larger molecules, e.g. fatty acids and cholesterol, has been known now for at least ten years (for review see Bloch, 1947, 1948; Popják, 1952), the reactive form of acetate in these biosyntheses, however, was not known until very recently. You will find frequently in the biochemical literature up to about 1950 the phrase 'acetate or a reactive derivative of it'. Nowadays this reactive derivative of acetate is referred to as acetyl-CoA. The enzymic formation of acetyl-CoA is an excellent illustration of the transfer of the chemical energy of ATP to other chemical groupings.

The identification of the so-called active acetate with acetyl-coenzyme A is largely due to the work of a group of distinguished biochemists in the U.S.A., including Lipmann, Ochoa, Stadt-

of the lactating goat, we were able to show that the synthesis proceeds in steps, in each step an acetate unit being added to the carboxyl-end of a fatty acid chain already synthesized (Popják, French, Hunter and Martin, 1951). These steps could be depicted as follows: the condensation of two acetate units gives butyric acid,



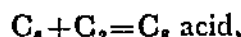
$$C_2 + C_2 = C_4 \text{ acid.}$$

The addition of another acetate unit to the carboxyl-end of butyric acid results in caproic (hexanoic) acid;



$$C_4 + C_2 = C_6 \text{ acid.}$$

The lengthening of the carbon-chain of caproic acid yields caprylic (octanoic) acid;



$$C_6 + C_2 = C_8 \text{ acid,}$$

and so on. There are good reasons, supported by experimental evidence, for believing that the fatty acids in organs other than the mammary gland are synthesized also in this stepwise manner (Popják, 1952). This view is also in harmony with the knowledge concerning the reactions of acetyl~S-CoA; the biosynthesis of fatty acids may then be represented in a general form to proceed according to the equation (10):



$$\text{acyl}_{(2n)}\sim\text{S-CoA} + \text{acetyl}\sim\text{S-CoA} = \text{acyl}_{(2n+2)}\sim\text{S-CoA} \quad (10)$$

in which the $\text{acyl}_{(2n)}$ denotes a fatty acid radical containing an even number of carbon atoms. The newly formed $\text{acyl}_{(2n+2)}$ fatty acid radical could be further elongated by reacting with another molecule of acetyl~S-CoA or could be removed from the reaction-chain by esterification with glycerol. The mechanism of fatty acid synthesis just outlined is still somewhat speculative since it has not been possible as yet to study the individual steps of fatty acid synthesis with enzymes isolated from animal cells. A modest beginning has been made recently towards this goal. Brady and Gurin (1952) and Miss Tietz and I (Popják and Tietz, 1953) have been able to obtain from pigeon-liver and mammary gland respectively particle-free enzyme preparations which, under appropriate conditions, synthesize fatty acids from acetate *in vitro*. These preparations are still crude and probably contain several enzymes in addition

This type of transfer of energy from ATP to another chemical grouping may be regarded as a model for chemical energy-transfer mechanisms resulting in an activation of substances for biosynthetic purposes. As I mentioned earlier, energy which is liberated in the course of degradation of foodstuffs as heat is lost entirely and cannot be utilized for syntheses by the cell. In the above set of reactions leading to the activation of acetate, it is important to notice that the energy is not liberated but is being handed along a chain of chemical reactions with very little loss, this being indicated by the fact that the individual reactions (a), (b) and (c) were all easily reversible (Jones, Lipmann, Hilz and Lynen, 1953), and that the free energy of the acetyl~S-CoA is between 10–12,000 calories (Stern, Ochoa and Lynen, 1952). Thus the acetyl~S-CoA complex contains now the potential energy required for synthetic reactions involving acetate units. For example, the free energy of formation of citrate from oxaloacetate and acetate is only + 4,680 calories (Kaplan, 1951), and it is probable that the value for the condensation of acetate units to fatty acids is not greatly different. Certainly acetyl~S-CoA is the form of 'active' acetate in the acetylation reactions of foreign amines (Kaplan and Lipmann, 1948), like sulphonamides and *p*-aminobenzoic acid, and in the biosynthesis of acetoacetic and of citric acid (Stern, Ochoa and Lynen, 1952).

It seems now probable that in the synthesis of fatty acids from acetate likewise, acetyl~S-CoA is the reactive substance. It has been shown with the aid of isotopic tracers that both carbon atoms of acetate are incorporated into fatty acids and that the entire carbon skeleton of these acids can be built up from acetate. For example, after an animal is given acetate labelled in the carboxyl-C with radioactive carbon, the fatty acids in the animal will contain the label in the odd-numbered carbon atoms ($R - CH_2\ C^*H_2\ CH_2\ C^*OOH$), whereas after the administration of acetate *labelled in the methyl carbon, the* even-numbered carbon atoms of the fatty acids will contain the isotope ($R - C^*H_2\ CH_2\ C^*H_2\ COOH$) (cf. Popják, Hunter and French, 1953). In an investigation into the mechanism of biosynthesis of fatty acids from acetate in the mammary gland

rabbits (Popják, unpublished observations), the two species in which such measurements were made. That this metabolite never accumulates in appreciable amounts is due to the fact that it is being used for the various bodily reactions as fast as it is being made; it is being oxidized in the citric acid cycle as a source of energy, it is used in acetylations and in various biosynthetic reactions, the synthesis of fatty acids being only one of these.

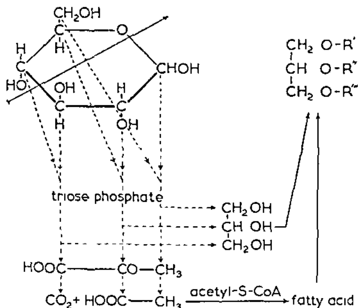


Fig. 2. Scheme of [illegible]

So far I have tried to build up a picture of enzymic synthesis through individual reactions and have illustrated the process by the example of formation of fat.

Some Methods in the Study of Biosynthesis and Regeneration of Body Constituents

It is befitting in this series of lectures, opened by Sir Henry Dale with an address on scientific methods in medical research,

to other substances. Our own preparations, which can be regarded essentially as the cell-sap of mammary gland cells diluted with buffer, certainly contain coenzyme A, diphosphopyridine-nucleotide and cytochrome in addition to enzymes; they will effect the synthesis of fatty acids from acetate only if ATP is added. Presumably the role of ATP in these mammary gland preparations is the activation of acetate with coenzyme A.[1]

Of course, the formation of fatty acids is only one facet of the many problems connected with the biosynthesis of fats and lipids, all of which are complex molecules; fatty acids in free form occur in tissues and body-fluids in minute amounts only. The conversion of carbohydrate, notably of glucose, into neutral (glyceride) fat is a good illustration of the processes of degradation and resynthesis occurring side by side in the body, to which I alluded at the beginning of this lecture. My colleagues and I have studied this problem in recent years with the aid of glucose labelled with ^{14}C (Popják, Hunter and French, 1953); the picture that emerged is summarized in Figure 3. The first stage of this process is the degradation of glucose in the glycolytic cycle which gives, through a common intermediate, glycerol on the one hand and pyruvic acid on the other. The oxidative decarboxylation of pyruvic acid leads to acetic acid, from which the fatty acids are then synthesized. The esterification of the fatty acids with the glycerol gives finally the whole molecule of a triglyceride fat. The magnitude and speed of these reactions may be assessed from our observations that in fully-fed normal rabbits all the glycerol in liver fat is newly formed from glucose in about six hours. With the exception of ruminant animals, there are usually only milligram quantities of acetic acid present in the various body tissues and fluids at any given time, yet this metabolite is formed continuously at the rate of about 1g./100g. body weight/day, both in rats (Bloch and Rittenberg, 1945), and in

acids was the same as in the nitrogen of air (99·62 per cent ^{14}N and 0·38 per cent ^{15}N) (Schoenheimer and Rittenberg, 1939). Or, again, the two naturally occurring isotopes of hydrogen are found in biological substances in the same proportion as in natural waters (99·98 per cent ^{1}H and 0·02 per cent ^{2}H). But if we give animals heavy water to drink, i.e. a water in which the concentration of deuterium (^{2}H) is greater than the natural 0·02 per cent, all the reactions of the body will take place in a medium containing heavy water. It was mentioned earlier that several of the biosynthetic reactions are reductive processes, e.g. the synthesis of fatty acids and of cholesterol from acetate. The hydrogen donors for these reductions are in rapid equilibrium with the hydrogen of body water and, therefore, in all such reductive syntheses deuterium will be built into the newly formed molecules (into the reduced chemical groupings) in the same ratio as it is present in the body water. In other words, the synthesized molecules will become labelled with deuterium. Deuterium was the first among the isotopes to be applied to the study of biological synthesis; the development of the technique is due entirely to Schoenheimer and Rittenberg at the Columbia University, New York (see Schoenheimer, 1941). If the concentration of heavy water in the body is kept constant (usually at 1–2 per cent level), then it is possible to calculate the rate of synthesis of several substances from the rate at which deuterium appears in these in a stably-bound form. Conversely, by measuring the rate at which the 'label' disappears, the rate of destruction of a substance can be measured. When these techniques were applied to the study of fat metabolism, it was found that in animals which kept their body weight constant the rate of synthesis of fatty acids was equal to the rate of their destruction or utilization (see Schoenheimer, 1941). This was one of the earliest observations providing clear-cut experimental evidence to show that the seeming constancy of the composition of the body is but the result of an equilibrium between synthesis and degradation, the two processes balancing each other. These observations have been extended to other body constituents also, to various lipoids, sterols, proteins, carbohydrates and other substances, and these, with the exception of a few, were

that I should also describe briefly some of the methods used in the study of biological synthesis and, parallel with this, should give an account of the balance of body constituents as this is affected by the processes of degradation and resynthesis mentioned in my introductory remarks.

There are two principal modern methods available for the study of biosynthesis: (*a*) the use of isotopes as tracers and (*b*) the use of isolated enzyme systems. Our present-day ideas concerning biological synthesis are very largely derived from the successful application of these two methods, separately or in combination. In earlier times, when such techniques were not available, the synthesis of a body constituent could be shown to occur only by careful and often tedious balance-sheet experiments demonstrating *a net increase in the body of a substance* not present in the diet. For example, it was by such a technique that the synthesis of fat from carbohydrate had been demonstrated in pigs about one hundred years ago by Lawes and Gilbert (1866). I would like to emphasize that a net increase in the amount of a particular substance had to be demonstrated before a conclusion could be reached from such experiments. It can be readily appreciated that experimentation of this kind often involved conditions far from physiological, and furthermore very little, if any, information could be gained as to the intermediary stages of the synthesis.

The discovery and isolation of the naturally occurring stable isotopes and then the production of radioactive isotopes of practically all elements of biological importance placed in the hands of biologists a most powerful tool of research, second in importance only to chromatography. It is possible to study the processes of biological synthesis in great detail with their aid under entirely physiological conditions. We have at our disposal isotopes of hydrogen, carbon, nitrogen, oxygen, phosphorus, sulphur, sodium, potassium, iodine, to name only a few. The principle of the application of these isotopes to biological research is quite simple and is based on the fact that, within certain limits, cells cannot distinguish between isotopes of the same element. This was shown, for example, by the observation that the ratio of the two stable isotopes of nitrogen in proteins and in amino

acids was the same as in the nitrogen of air (99·62 per cent ^{14}N and 0·38 per cent ^{15}N) (Schoenheimer and Rittenberg, 1939). Or, again, the two naturally occurring isotopes of hydrogen are found in biological substances in the same proportion as in natural waters (99·98 per cent ^{1}H and 0·02 per cent ^{2}H). But if we give animals heavy water to drink, i.e. a water in which the concentration of deuterium (^{2}H) is greater than the natural 0·02 per cent, all the reactions of the body will take place in a medium containing heavy water. It was mentioned earlier that several of the biosynthetic reactions are reductive processes, e.g. the synthesis of fatty acids and of cholesterol from acetate. The hydrogen donors for these reductions are in rapid equilibrium with the hydrogen of body water and, therefore, in all such reductive syntheses deuterium will be built into the newly formed molecules (into the reduced chemical groupings) in the same ratio as it is present in the body water. In other words, the synthesized molecules will become labelled with deuterium. Deuterium was the first among the isotopes to be applied to the study of biological synthesis; the development of the technique is due entirely to Schoenheimer and Rittenberg at the Columbia University, New York (see Schoenheimer, 1941). If the concentration of heavy water in the body is kept constant (usually at 1–2 per cent level), then it is possible to calculate the rate of synthesis of several substances from the rate at which deuterium appears in these in a stably-bound form. Conversely, by measuring the rate at which the 'label' disappears, the rate of destruction of a substance can be measured. When these techniques were applied to the study of fat metabolism, it was found that in animals which kept their body weight constant the rate of synthesis of fatty acids was equal to the rate of their destruction or utilization (see Schoenheimer, 1941). This was one of the earliest observations providing clear-cut experimental evidence to show that the seeming constancy of the composition of the body is but the result of an equilibrium between synthesis and degradation, the two processes balancing each other. These observations have been extended to other body constituents also, to various lipoids, sterols, proteins, carbohydrates and other substances, and these, with the exception of a few, were

found to behave similarly. Although Sir Frederick Gowland Hopkins described the life of the cell as early as 1913 in his address to the British Association in Birmingham as 'a dynamic equilibrium in a polyphasic system', such a view at that time was more prophetic than based on experimental fact. The study of animal metabolism, in particular of biosynthetic phenomena, with isotopic tracers gave, however, solid backing to these visionary ideas which were fully expounded by Schoenheimer (1941) in his monograph 'The dynamic state of body constituents'.

Of course, not all biological systems are in a steady state. The most obvious example is the growing animal both in the prenatal and postnatal period, or a growing tumour. We may pose the question, what is the underlying principle of growth? Is it associated with an increased rate of synthesis or with an inhibition of degradation? When isotopic tracer technique was applied to such problems a fairly clear-cut answer was obtained, at least as far as some tissue constituents are concerned. When pregnant rabbits were given heavy water to drink, it was found that the incorporation of deuterium into foetal cholesterol, for example, was faster than into cholesterol in either the maternal liver or in the placenta (Popják and Beeckmans, 1950). It was thus possible to conclude that foetal tissues synthesize cholesterol from small molecules. However, when the rates of synthesis and degradation were calculated from a comparison of the isotope data with growth rate curves it became apparent that in foetal tissues the rate of synthesis of cholesterol is about the same as in maternal liver cells (about 50–60 mg./100 g. tissue/day), but degradation is very much slower or completely absent resulting in a net accumulation of the substance. Similar conclusions were reached in respect of foetal fats, too. Rittenberg, Sproul and Shemin (1948) also concluded that the growth of the regener- [illegible] ...ectomy is the result of the inhibi- [illegible] ...rotein synthesis in the regenerating liver [illegible] faster than in normal animals. The conclusion that growth results not so much from an increased rate of synthesis as from an absence of, or decreased rate of, degradation appears to be reasonable on the

grounds of reaction kinetics also, since generally it is easier to slow down or stop a biological reaction than to speed it up.

The investigations with heavy water just quoted were examples of an experimental design, in which the concentration of an isotope, in an appropriate chemical form, is maintained constant in the body. Such experiments lend themselves very readily to the study of biosynthetic processes in a quantitative manner.

In another type of experiment the isotopically labelled substance, e.g. an amino acid labelled with ^{15}N or with ^{14}C, is administered for a short period only, either in a single dose or in divided doses in the course of a few hours or few days. Body constituents, which are synthesized from such labelled amino acids and which are being continuously regenerated, will show the kind of isotope incorporation illustrated in Figure 4. The curves in the top part of this figure represent the change in ^{15}N-content of normal plasma proteins and of antibody in a rabbit actively immunized with type III pneumococcus and fed ^{15}N-labelled glycine for three days (Schoenheimer, 1941). The isotope concentration in these proteins rises during the administration of the labelled precursor, then reaches a maximum, after which it declines exponentially. The declining part of the curve need not necessarily be, and seldom is, represented by a single exponential equation, but may be a combination of two or more exponentials. The experiment shown in Figure 4 demonstrated an interesting phenomenon, viz. that not only normal plasma proteins were being continuously resynthesized, but also antibodies produced by active immunization, even though the total antibody titre in the serum was falling. Presumably in this instance the rate of degradation or elimination of immune body protein was faster than its synthesis, hence the decreasing titre in the serum. In contrast with this, in passive immunity the antibodies injected into an animal show no evidence of regeneration, i.e. no label appears in the injected antibody when a labelled amino acid is given simultaneously. These basic informations on the behaviour of antibodies have been confirmed and extended recently by McFarlane and Humphrey (1954) with the aid of ^{14}C-labelled amino acids.

The magnitude of antibody synthesis may occasionally greatly exceed the synthesis of normal plasma proteins. McFarlane and Humphrey (1954) calculated for example that one of their hyperimmunized rabbits produced antibody four times as fast

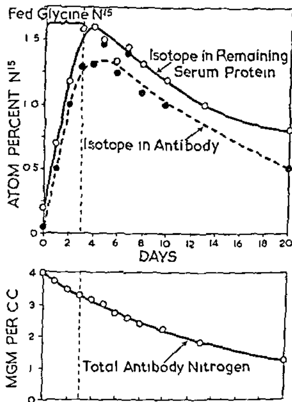


FIG. 4 *Upper part.* ^{15}N-content of normal plasma proteins and of antibody in a rabbit actively immunized with type III pneumococcus and fed ^{15}N-labelled glycine for three days. *Lower part·* Total antibody nitrogen in plasma during the experiment (Schoenheimer, 1941)

as it did normal serum globulin and that during a 19-day period of observation this animal synthesized about 18 g. of antibody in response to an antigenic stimulus of 2·5 mg. of pneumococci of which not more than 10 per cent was antigenic capsular polysaccharide.

Among the exceptions to the general rule of continuous syn-

thesis and degradation of body constituents, the most notable is the haemoglobin of the mammalian red cell. Shemin and Rittenberg (1945, 1946 *a*, *b*) made the significant discovery that the nitrogenous precursor of the protoporphyrin of haemoglobin was the amino acid glycine. This discovery has led to a considerable elucidation by the Columbia workers and by Muir and Neuberger (1949, 1950) in this country, of the biosynthesis of protoporphyrin. Shemin took by mouth 66 g. of ^{15}N-glycine in the course of three days and then they observed the concentration of ^{15}N in the haemin prepared from his red corpuscles. It was found that the ^{15}N-concentration in the haemin rose steeply, reaching a maximum value about 25 days after the labelled glycine was taken, then remained almost constant for another 70 days, and then fell along an S-shaped curve. The shape of this curve is totally different from that found for body constituents which continuously regenerate. From the shape of the isotope concentration/time curve of haemin it appeared that the probability that a haemoglobin molecule may be degraded is a function of its age. Since the principle of 'ageing' cannot be applied to molecules another explanation had to be sought to account for this particular type of curve. If it be assumed that the haemoglobin molecule, laid down during maturation of the erythrocyte, is not involved in the flux of synthesis and degradation like other substances in the body, but remains in the red cell until the cell disintegrates or dies, the shape of the isotope concentration/time curve of haemin observed by Shemin and Rittenberg can be adequately explained. This assumption has been generally accepted as covering the facts. The observation afforded also a new method for measurement of the mean life span of human erythrocytes and gave in the subject studied a value of 127 days. The mean life of the erythrocytes had previously been estimated by following the survival of blood-group O cells injected into a blood-group A recipient by the agglutination technique of Ashby (Callender, Powell and Witts, 1945). This method yielded a value of 120 days. A similar value was found by Jope (1946) who measured the disappearance of sulphaemoglobin from the blood of men who had been exposed to trinitrotoluene. Both these non-isotopic techniques are, in

principle, labelling methods, and they assume that the life-span of the labelled cells is identical with that of normal cells, an assumption which may not necessarily be correct. This criticism does not apply to the isotopic tracer technique, the advantage of which is that the life of the individual's own unaltered red cells is being measured.

The Columbia workers were quick to realize the possibilities their discovery offered, and extended their investigations to the study of haem synthesis and red cell dynamics in pathological conditions such as untreated and treated pernicious anaemia, polycythaemia vera, and sickle-cell anaemia (London, Shemin, West and Rittenberg, 1949). Thus they were able to show that *polycythaemia vera, in its fully developed form, is characterized* by an abnormally high rate of haematopoiesis and a normal life-span of the red cells. In sickle-cell anaemia the isotope-concentration in the haemin rose rapidly to a peak on the seventh day after the administration of the ^{15}N-labelled glycine, and immediately began to fall. The analysis of the curve led to the conclusion that in sickle-cell anaemia the red cells are destroyed at random, independently of their age, and therefore their survival cannot be described properly by a 'mean-life'. It was found at the same time that the rate of haemoglobin synthesis was about three times as fast as normal.

I am deeply conscious of the fact that the picture I have given of biological synthesis is deficient and inadequate in many respects. However, the subject matter provides my apologia in that the problem of biosynthesis of any one cell constituent could have provided enough material for this lecture or more. I have attempted to describe biosynthesis as a process integrated with the energy metabolism of the cell and to stress principles rather than detail, and I have endeavoured to show that the study of these phenomena has an intrinsic value, not only to the biochemist for the understanding of fundamental processes of life, but also to the clinician for the understanding of human disease.

REFERENCES

BALL, E. G (1944). *Ann. N.Y. Acad. Sci* **45**, 363.

BLOCH, K. (1947). *Physiol. Rev.* **27**, 574.

BLOCH, K. (1948). *Cold. Spr. Harb. Symp. quant. Biol* **13**, 29

BLOCH, K. and RITTENBERG, D. (1945). *J. biol. Chem.* **159**, 45.

BRADY, R. O. and GURIN, S (1952). *J. biol. Chem.* **199**, 421.

BURTON, K. and WILSON, T. H. (1953). *Biochem J.* **54**, 86

CALLENDER, S. T. E., POWELL, E. O. and WITTS, L. J. (1945). *J. Path. Bact.* **57**, 129.

[illegible]

GAFFRON, H (1935). *Biochem. Ztschr* **275**, 301.

GAFFRON, H. (1946) In *Currents in Biochemical Research*, edited by D. E. Green, Interscience Publishers, New York.

HILL, R. and SCARISBRICK, R. (1940) *Nature, Lond.* **146**, 61.

JONES, M. E., LIPMANN, F, HILZ, H and LYNEN, F (1953) *J. Amer. chem. Soc.* **75**, 3285

JOPE, E M. (1946). *Brit. J indust. Med.* **3**, 136

KAPLAN, N. O (1951). In *The Enzymes*, edited by J. B. Sumner and K. Myrback, vol. 2, p 55 New York, 1951

KAPLAN, N. O and LIPMANN, F. (1948). *J biol. Chem* **174**, 37.

KREBS, H. A. (1953). *Brit med. Bull* **9**, 97.

LAWES, J B. and GILBERT, J H (1866). *Phil Mag, London*, 4th series, **32**, 439

LEHNINGER, A L. (1949) *J. biol Chem* **178**, 625.

LEHNINGER, A L (1951) *J. biol Chem* **190**, 345

LEHNINGER, A L and SMITH, S. W. (1949) *J biol Chem.* **181**, 415

[illegible] Congrès [illegible]

LONDON, I. M, SHEMIN, D., WEST, R and RITTENBERG, D (1949) *J. biol. Chem* **179**, 463.

LYNEN, F, REICHERT, E and RUEFF, L (1951) *Ann Chem.* **574**, 1

McFARLANE, A S and HUMPHREY, J (1954). *Biochem. J* **57**, 186

[illegible]

[illegible] P (1951). *Biochem J* **48**, 612

POPJÁK, G., HUNTER, G D. and FRENCH, T. H (1953) *Biochem. J* **54**, 238.

POPJÁK, G. and TIETZ, A. (1953). *Biochim. Biophys. Acta*, **11**, 587.
PORTER, H. K. and MARTIN, R. V. (1952). *J. exp. Bot.* **3**, 326.
RITTENBERG, D., SPROUL, E. E. and SHEMIN, D. (1948). *Federation Proc.* **7**, 180.
SCHOENHEIMER, R. (1941). *The dynamic state of body constituents*, Cambridge, Mass., 2nd ed., 1946.
[illegible] 7, 285.
[illegible]
STERN, J. R., OCHOA, S. and LYNEN, F. (1952). *J. biol. Chem.* **198**, 313.

III

The Genetics of Some Biochemical Abnormalities

J. B. S. HALDANE

I HAVE chosen this title deliberately, rather than such a title as 'The Genetics of some Metabolic Diseases', for two reasons. A genetically determined or inborn biochemical abnormality may be a necessary, but not sufficient, condition for a disease in men, animals or plants. The disease will develop if you have the metabolic abnormality *and* an infection or some other environmental factor which may be anything from overeating to exposure to sunlight. In many cases, so far as we know, a biochemical abnormality is completely harmless. However, if it is rare there is always the presumption that it has not spread because it is in some circumstances harmful. Secondly a biochemical abnormality may mean, not that there is any predisposition to a disease, but that if two people with the abnormality have children, these children are liable to the disease. As we shall see, abnormalities of this type may be advantageous.

I shall speak largely about fairly rare conditions for a very simple reason. If a gene is fairly common it cannot (except in one set of cases of which I shall speak later) be very harmful. If it reduced the chances of its carriers of having children appreciably it would soon become rare, though it might be kept in being by mutation.

Garrod, the founder of human biochemical genetics, realized that these abnormalities were conditions determined by abnormal genes. He did not speculate on how the genes determine the abnormality. This we can now do, thanks to the extraordinarily

simple results of human genetical serology. We know that in almost all cases there is a very simple relation between a gene and an antigen. Everyone who has the A agglutinogen on his or her corpuscles has at least one A gene per nucleus which he or she received from a parent, and will transmit it to about half his or her children. Similarly everyone with such a gene has the antigen on his or her corpuscles. The simple relation is not true for small molecules. For example purple sweet-pea flowers contain a glycoside of malvidin, but at least four different genes must be present before this substance is formed. According to which is missing one gets crimson, salmon or white petals. It is, therefore, tempting to regard antigens as primary products of genes, and to take the same view as regards many of the enzymes concerned in intermediary metabolism. It is certainly not true of all proteins. Filitti-Wurmser (1954) and her colleagues have shown that the anti-B agglutinogen of A_1O individuals is different from that of A_1A_1 and OO individuals, and is not formed by mixing them *in vitro*.

Perhaps the most fully investigated case in human biochemical genetics is that of sickle cell anaemia or drepanocytosis. I refuse to use the American word 'sicklemia'. Normal people have haemoglobin which is homogeneous in an electric field. So do their children Some negroes have a mixture of two haemoglobins which differ in their isoelectric points both when reduced and when combined with carbon monoxide. These people are heterozygotes. If they marry normal people about half the children have normal haemoglobin and half mixed haemoglobin. They have one gene for making normal haemoglobin and one for making abnormal haemoglobin. When two such heterozygotes marry, a quarter of their children get the abnormal gene from each parent, and have nothing but abnormal haemoglobin. These people are very seriously handicapped, because the abnormal haemoglobin is insoluble when reduced. In the venous blood their corpuscles assume odd shapes, and these corpuscles have a short life. In consequence they suffer from a severe haemolytic anaemia usually fatal in childhood Men and women with mixed haemoglobin seem to be normal. We shall see that in one situation they are super-

normal. The difference between the haemoglobins is certainly in the globin part of the molecule.

The main facts concerning the genetics were established by Neel, those concerning the chemistry by Pauling, Itano, Singer and Wells. I do not propose to give the references, since a very full bibliography of this and most of the other topics on which I am speaking today is given in Harris' *An Introduction to Human Biochemical Genetics*: I shall only give references not found there or in my *The Biochemistry of Genetics*.

Three further points complicate the situation. Many, but not all, of the anaemics, besides sickle cell haemoglobin, have anything up to 25 per cent of foetal haemoglobin. This can be thought of as a physiological compensation. The percentage of insoluble haemoglobin in 32 heterozygotes varied from 45 per cent to 22 per cent. Finally, other genes have been found which produce still other types of haemoglobin, and may cause anaemia when corpuscles contain a mixture of one of these with sickle cell haemoglobin.

This condition is remarkably common. It is said that 9 per cent of American negroes are heterozygotes. If so one negro baby in 500 is anaemic for this reason, and will probably die before puberty.

What is worse, in tropical Africa it is still commoner. 18 per cent of babies in many populations are heterozygotes, which means that about 1 per cent of the children are anaemic. Professor Gates suggested that the anaemia only arose after crossing with Europeans. There is no evidence for this hypothesis, which would have pleased Hitler, on the contrary about 75 per cent of anaemics in Africa die before they become adult, but the question certainly arises 'Why has not natural selection extinguished the gene?' The answer was found by Allison (1954). He finds that the frequency of the anaemia is correlated with that of malignant tertian malaria, the frequency of heterozygotes occasionally reaching 40 per cent. In Uganda the heterozygotes are substantially immune to malignant tertian malaria. Of 15 inoculated, only 2 developed mild infections even after a second inoculation. Of 15 controls, 14 were infected. Perhaps *Plasmodium falciparum* finds sickle cell haemoglobin indigestible.

simple results of human genetical serology. We know that in almost all cases there is a very simple relation between a gene and an antigen. Everyone who has the A agglutinogen on his or her corpuscles has at least one A gene per nucleus which he or she received from a parent, and will transmit it to about half his or her children. Similarly everyone with such a gene has the antigen on his or her corpuscles. The simple relation is not true for small molecules. For example purple sweet-pea flowers contain a glycoside of malvidin, but at least four different genes must be present before this substance is formed. According to which is missing one gets crimson, salmon or white petals. It is, therefore, tempting to regard antigens as primary products of genes, and to take the same view as regards many of the enzymes concerned in intermediary metabolism. It is certainly not true of all proteins. Filitti-Wurmser (1954) and her colleagues have shown that the anti-B agglutinogen of A_1O individuals is different from that of A_1A_1 and OO individuals, and is not formed by mixing them *in vitro*.

Perhaps the most fully investigated case in human biochemical genetics is that of sickle cell anaemia or drepanocytosis. I refuse to use the American word 'sicklemia'. Normal people have haemoglobin which is homogeneous in an electric field. So do their children. Some negroes have a mixture of two haemoglobins which differ in their isoelectric points both when reduced and when combined with carbon monoxide. These people are heterozygotes. If they marry normal people about half the children have normal haemoglobin and half mixed haemoglobin. They have one gene for making normal haemoglobin and one for making abnormal haemoglobin When two such heterozygotes marry, a quarter of their children get the abnormal gene from each parent, and have nothing but abnormal haemoglobin. These people are very seriously handicapped, because the abnormal haemoglobin is insoluble when reduced. In the venous blood their corpuscles assume odd shapes, and these corpuscles have a short life. In consequence they suffer from a severe haemolytic anaemia usually fatal in childhood. Men and women with mixed haemoglobin seem to be normal We shall see that in one situation they are super-

with 100 per cent methaemoglobin. Such a child would of course die some time before birth.

The most typical biochemical abnormalities, as Garrod recognized fifty years ago, are metabolic blocks. They are due, at least in some cases, to the absence of the enzyme catalysing a particular process. Let me take an example with some comic features. The catalase of blood is well known. A drop of blood makes a hydrogen peroxide solution fizz merrily. But if I ask you its function, I may get a variety of answers. No answer given *before 1952 was correct, because no one had observed a mammal*, let alone a man, without catalase. In fact it has the function of protecting the teeth.

Takahara (1952) described nine cases, in three sibships all from marriages of normal cousins, of progressive oral gangrene. The genetics suggest strongly, but do not prove conclusively, that it is a recessive. The gums and cheeks are ulcerated. All teeth may be lost at age of 15. On recommending hydrogen peroxide as a mouth wash he was astonished to find no bubbles, and blackened ulcers. When blood from these people is added to hydrogen peroxide it blackens and does not bubble. The catalase content is less than one-thousandth of normal; but peroxidase is present. This suggests that normally blood catalase is a defence against haemolytic streptococci. The Editor of the *Lancet* suggests that cases may be found in Europe.

Where do we go from here? The obvious point is to search for similar cases. Less obvious is to look for low catalase in milder cases of oral sepsis It is at least conceivable that people may be found in whom blood catalase is low but not absent, and that they may be particularly liable to attacks by haemolytic streptococci. I think it is probable that we could go a lot further. Most rabbits have an atropinesterase in their serum The absence of it is recessive. But Sawin and Glick found that heterozygous rabbits only have about half as much of this esterase as normal ones. Presumably the dominant genes make esterase independently of one another.

It might therefore be possible, by doing quantitative measurements of blood catalase, to pick out heterozygotes Clearly two such heterozygotes should not marry. There is still

A simple calculation shows that, if all the anaemics died, the heterozygotes would have to have an advantage of about 11 per cent over normals to keep a population in equilibrium with 81 per cent normal, 18 per cent heterozygous, and 1 per cent anaemic, if all the latter died before reproducing. In some environments the advantage of heterozygotes should be greater than 20 per cent.

As Professor Penrose has pointed out to me in conversation, Allison's observations are doubly paradoxical from the point of view of widely held theories about congenital disease. On the one hand an infectious disease has been responsible for the spread of a congenital disease in no obvious way related to it. On the other, chemotherapy, so far from causing racial degeneration, will presumably lead to the disappearance of this disease.

I give this case in some detail, because it illustrates the possibilities of biochemical eugenics. On the face of it, sickle cell anaemia is a recessive disease; and since almost all the parents are normal, no obvious eugenic measures can be taken against it. In fact one can pick out heterozygotes; and any negro, before marriage, would be well advised to see that he and his intended bride were not both heterozygotes. This would stop about 1 per cent of negro marriages in the U.S.A. and 3 or 4 per cent in Central Africa. It would abolish sickle cell anaemia. Can we hope to do the same with a variety of rare conditions found in this country which, however, among them account for a good deal of chronic disease? Not yet, but perhaps before A.D. 2000.

Let me take an example of a so-called dominant condition with a very similar etiology. Horlein and Weber studied a family in which methaemoglobinaemia had been handed down for at least four generations About 20 per cent of the haemoglobin was present as methaemoglobin, and they showed that it had an abnormal spectrum and was not reduced by ascorbic acid or methylene blue, as in other types. The abnormality was shown to be in the globin part of the molecule, as in sickle cell anaemia. It seems likely that if two methaemoglobinaemic members of this family married, they might produce a child

Metabolic studies on man have shown that some amino-acids such as phenylalanine and tryptophan are essential. They must be present in the diet. Others, such as tyrosine and serine, are not essential. They can be made from other amino-acids. Further, one of the commonest kinds of biochemical mutant in moulds and bacteria is one which requires some particular constituent, say lysine or aneurine, in the culture medium, which the normal form can make for itself.

Phenylketonuria seems to be due to a mutation of this type in man. If a normal human being is fed with phenylalanine, tyrosine rapidly appears in the plasma. It does not in a phenylketonuric, though experiments with 'labelled' atoms show that a little tyrosine is slowly formed. Thus tyrosine is presumably an essential amino-acid for such people. The metabolic step which is blocked is the conversion of phenylalanine to tyrosine. Whether the mental defect is due to chronic intoxication with phenylalanine or one of its catabolic products is not yet known with certainty. It should be known within a year or so, as attempts are being made to keep several phenylketonurics on a diet in which phenylalanine is reduced to a minimum.

Phenylketonuria has a variety of effects. It was discovered because it is responsible for about 1 per cent of the severe congenital mental defects (imbecility and idiocy) in this country. However, not quite all phenylketonurics are certifiable. Some are merely backward, though I doubt if any have reached an I.Q. above 80. Further, Penrose found that on the whole they have smaller heads and lighter hair than the rest of the population. Melanins are of course formed from aromatic amino-acids. The term phenylpyruvic amentia is much less satisfactory than phenylketonuria because not all phenylketonurics are aments, and it is impossible to draw a sharp line between those who are mentally defective and those who are merely stupid. On the other hand the chemical abnormality is quite sharply defined. I shall give other examples later where a biochemical abnormality may be symptomless.

We know that phenylalanine is oxidized to tyrosine under the influence of an enzyme found in human and other mammalian livers, but not in a number of other organs. But we do not

another possibility. It is possible that the recessive genes, instead of making catalase, make some other protein which differs from it antigenically, and is therefore serologically detectable.

Let me take another case, which illustrates two points. In a Northern Irish family methaemoglobinaemia was found in five brothers and sisters out of nine, and in no other relatives. This looks like a recessive. In this family the haemoglobin was apparently of normal structure, but Gibson and Harrison found that a flavoprotein reducing enzyme, 'co-enzyme factor 1', was missing from the corpuscles. Treatment with ascorbic acid or methylene blue was quite effective in reducing the haemoglobin. You see that in this family and Horlein and Weber's family two different genes were producing the same effect by quite different methods. Both the genetics and the therapeutics were different. One may compare these two methaemoglobinaemias with bacillary and amoebic dysentery. In the Irish family it might be quite possible to pick out heterozygotes by quantitative measurements on corpuscular enzymes.

Let us turn to a much more serious and unfortunately commoner condition, phenylketonuria. Phenylketonurics excrete phenylpyruvic acid, phenylalanine, phenyl-lactic acid, and other related substances in their urine in quantities of the order of a gramme per day. They are usually mentally defective. There is no doubt that the character is recessive. Phenylketonurics very rarely have children. All the 94 parents of the 85 described by Munro were normal If a character is recessive it [illegible] heterozy- [illegible] when [illegible] ch include at least one recessive, the frequency is well above a quarter. Clearly it is 100 per cent in sibships of one. It should be 57 per cent in sibships of two, falling to 33 per cent in sibships of five and 26·7 per cent in sibships of ten. The numbers found in Munro's 47 sibships were 85 out of 226, the expectation being $77{\cdot}4 \pm 5{\cdot}4$. Again five of these 47 sibships, or 11 per cent, were derived from first-cousin marriages, and others from more distant relatives.

a generalized amino-aciduria, but no symptoms. They are taking dietary precautions, and if they develop symptoms like their sister, these should be readily curable.

Finally cystinuria is a sympton of Wilson's disease, again probably recessive, but with severe involvement of the liver and nervous system. Here it is more doubtful if the kidneys are involved, and the recent work of Schönberg and Gitlin (1952) suggests that the primary defect may be a failure to form caeruloplasmin, a copper-containing alpha globulin which is an oxidase. Turpin, Jerome and Schmitt (1953) confirm this finding. But they also find low values for caeruloplasmin among the clinically normal relatives of children dying of hepatolenticular degeneration. It is too early to say whether these were symptomless cases of the abnormality or heterozygotes. The latter seems more probable.

I shall, I hope, be forgiven if I now speculate briefly on the future of medicine. As infectious and deficiency diseases are conquered, congenital diseases will become more important. We can classify many of them roughly as determined by dominant, sex-linked recessive, and autosomal recessive genes. Now a dominant gene rarely produces severe presenile symptoms in all its carriers. If it did it would not persist for long. Natural selection would soon wipe it out. If then it does not invariably produce symptoms it is theoretically possible to prevent it from ever doing so. Such an attempt should certainly be made in parallel with any attempt to eradicate the gene by eugenic measures. The recessive conditions are more tragic. Who does not know healthy parents who have borne one or more grossly abnormal children?

I wish to suggest that within a century it should be possible to detect heterozygosity for most of the more serious autosomal recessive conditions, and that an examination for them will be made at puberty as a routine. Let us see what this would mean. About 1 per cent of each sex is heterozygous for phenylketonuria, and quite possibly this is advantageous. If every person

know[1] whether this enzyme is lacking in phenylketonurics, or whether, perhaps, its action is blocked by some abnormal metabolic product. Nevertheless I hope to live long enough not merely to know the answer to these questions, but long enough to see a method of detecting heterozygotes. It is conceivable, for example, that the rate at which plasma tyrosine rises after [illegible] frequency of one phenylpyruvic in 40,000 with a few extra ones due to inbreeding.

Let me illustrate my theme with one more set of examples from work partly done by Harris in the department of which I am head at University College, though in fact my colleague Penrose is responsible for its human work. Garrod wrote of cystinuria as a metabolic abnormality. We now know that it is in fact a group of abnormalities due to different genes and causing different symptoms. One or more, probably two, different recessive genes cause the urinary excretion of cystine, lysine and arginine in quantities of the order of a gramme daily. Renal calculi may occur, but do not do so invariably, and no other symptoms are known. The error may be purely renal. But cystine crystals have been reported in the cornea and elsewhere. It is not inconceivable that tissues in general may have some difficulty in dealing with cystine, even if it is only in the kidneys that this causes serious trouble. I have dealt with this question elsewhere (Haldane, 1954).

In Fanconi's syndrome (a recessive abnormality of childhood) the urine contains not only cystine, arginine and lysine, but Uncle Tom Cobley and all, that is to say most of the amino-acids, glucose, and extra phosphate. The rickets which are its most serious symptom seem to be due to phosphate loss.

Dent and Harris described a pedigree of this syndrome which only appeared in adult life. The condition is probably recessive. One patient was bedridden with 'rheumatic pains', spontaneous

[1] *Jervis (1953)* found no phenylalanine oxidase at autopsy in the livers of two phenylketonurics, while demonstrating it in three controls. Probably, therefore, this enzyme is produced by a gene present in normal human beings, but not in phenylketonurics

TAKAHARA, S. (1952). Progressive oral gangrene probably due to lack of catalase in the blood (acatalasaemia). Report of nine cases. *Lancet*, **ii**, 1101.

TURPIN, R., JEROME, H. and SCHMITT, H. (1953). Study of variations of the caeruloplasmin by an easy technique. *Proc. Roy. Soc. Med.* **46**, 1061.

who was heterozygous knew the fact, one intended marriage in 10,000 would be contra-indicated. If a hundred such conditions could be detected, 1 per cent of marriages would be contra-indicated. I do not believe that it would even be necessary to prohibit them legally. No one wants defective children, and provided the carrying of such genes was regarded as a family joke like having a piebald uncle, rather than a guilty secret like having an uncle in Dartmoor, I think the possession of a common recessive would come out long before a couple fell deeply in love. In fact I can imagine a swain of A.D. 2054 telling his dancing partner 'By the way, I'm heterozygous for infantile amaurotic idiocy and porphyria: I hope you aren't', and getting his face smacked for sauciness.

This may sound ridiculous. But a physician of 1854 would have found at least equally ridiculous the notion that everyone should know the blood group of which they are a member, which, I think it will be agreed, is advisable. I must apologize if I have been too speculative. But if a layman is asked to address a medical audience, it must be expected that he will hesitate to speak on what is immediately practicable, but will attempt to speculate on how his work can be applied to medicine in the future.

REFERENCES

ALLISON, A. C. (1954) Protection afforded by the sickle cell trait against subtertian malaria infection. *Brit. med. J* 290.

FILITTI-WURMSER, S., JACQUOT-ARMAND, Y., AUBEL-LESURE, G and WURMSER, R (1954). Physico-chemical study of human isohaemagglutination. *Ann Eugen* 18, 183.

HALDANE, J. B. S. (1954). *The Biochemistry of Genetics* Allen and Unwin, London.

HARRIS, H. (1953) *An introduction to human biochemical genetics.* Galton [illegible]

JE [illegible]

Sc [illegible]

116, 484.

ings in photographs and drawings made, with exemplary patience, from fixed preparations collected at regular intervals after wounding by accident in human beings or deliberately, for experimental purposes, in animals. The introduction of the Clark-Sandison chamber technique now allows this process to be watched and photographed on cine films hour by hour for weeks on end. Fibroblasts can be traced to pre-existing tissue, and are seen to grow at the rate of about 0 2 mm. a day into the blood clot of the chamber, lagging behind the macrophages and at about the same time as the new blood vessels. In about six days the first fibres appear around fibroblasts close to the periphery of the chamber and they can be traced towards the centre at about the same rate as the fibroblasts. In due course there is formed a mature fibrous tissue of good quality, in which cells are not so numerous as formerly, have the characteristics of [illegible]
or [illegible]
[illegible]
satisfactorily in the living preparation and we must return to fixed preparations for the details. Fortunately we have ways of speeding up fibril formation by making animals, especially guinea-pigs, scorbutic and then supplying them with vitamin C. In scurvy the manufacture of intercellular materials is held up, but after the administration of antiscorbutics, appreciable amounts of intercellular substance are formed within 24 hours and in a few days this material is completely differentiated. We owe to Professor S. B. Wolbach (1933) of Harvard an excellent account of this process of repair, the injury taking the form of a haematoma produced in scorbutic guinea-pigs by excision of small pieces of extensor muscle. Repair begins at once by the migration of fibroblasts from adjacent tissues into the clotted haematoma and these cells continue to divide as they invade. But they form no fibrils and they remain separated from each other, retaining their embryonic shape. Capillaries, too, do not penetrate the haematoma for any considerable distance and often they persist as closed columns of endothelial cells. Soon a curious vacuolar material appears between the fibroblasts which is delicately acidophilic and often granular. The vacuoles appar-

IV

Tissue Repair

G. R. CAMERON

It is customary to introduce the study of regeneration by discussing the healing of skin wounds and, indeed, there can be no better mode of setting the scene, complex though the action be. I shall follow the time-honoured precedent by reminding you of the essential features of *healing by first intention* and by granulation and pass rapidly to some of the lessons we have learned in this way.

Skin Wounds

When an aseptic incision is made through the skin and subcutaneous tissue as in the course of a surgical operation, many cells along the plane of cutting are damaged or killed and many vessels are severed. From this injury comes, as the result of liberation of histamine, peptides and perhaps other irritant chemicals, a plastic inflammation which gives oedema and fibrin formation across the cut plane with leukocytic and macrophage migration from adjacent vessels to the vicinity of dead tissue. We believe that this inflammatory reaction to the severed tissue serves the purpose of removing dead material and blood clot as a preliminary to repair.

Within 24 hours of injury, normal connective tissue cells alongside the incision swell up, assume round or oval shapes and divide to give new generations of young connective tissue cells or fibroblasts. These in turn may divide again and again by mitosis but sooner or later become quiescent, elongate, take on spindle shapes and now are seen to be the centre of fibril formation. For many years histologists have depicted these happen-

cess I shall have more to say when discussing tendon repair and the second process will be encountered again in the more appropriate place of liver regeneration.

While the fibrous tissue cells are going through their cycles of division and fibrillation, new blood vessels are being modelled from solid endothelial buds with fine terminal prolongations which grow out from pre-existing vessels. The modern investigations of the Clarks (1932, 1933, 1936) and of Florey and his co-workers (Ebert and Florey, 1939; Ebert *et al.*, 1940) with the Clark-Sandison chamber inserted in rabbits' ears and more recently of Algire (1943) using a transparent chamber attached to the mouse's abdominal wall, and of Speidel with his ingenious technique for the tadpole's tail (see 1948) have confirmed the original discoveries of Thoma (Cameron 1952, for discussion) and taught us something about factors therein concerned. New blood vessels grow in the Clark-Sandison chamber at the rate of 0·2–0 6 mm. a day; rate of growth can be modified by temperature and the obstacles to be overcome, especially the consistency of blood clot which may lie in the way of the growing buds. Nuclei increase by mitotic division, blood flows into the lumina that soon form, sprouts join up to form at first loops and later complicated plexuses, while differentiation into [illegible]

the loops from parent vessels or quite likely they represent modified fibroblasts. Equally puzzling are the factors that bring about the orientation of these cells in the vessel walls and attract adventitial cells of reticulo-endothelial origin to the outer limits of the differentiating vessels. And here, too, we encounter forces by which the vasculature may be modelled and remodelled. The lumina of many of the capillaries can be obliterated, the solid cord of endothelium thus produced dividing and the ends retracting to the level of the next patent vessel. In other words, reversibility is a property associated both with the production of fibrils and new blood vessels.

Lymphatics grow in much the same way as the blood capillaries, emerging as solid sprouts from pre-existing lymphatics

ently are filled with fluid and some can be traced to fibroblasts as if they had been discharged from these cells. Wolbach maintains that this is not oedema because the vacuoles occur in connection with cells remote from blood vessels. More likely, they are cellular in origin, either degenerative or secretory. I would like to suggest that they come from mitochondria, but for this view there is no evidence as yet.

Now, if orange juice rich in vitamin C be administered daily to these guinea-pigs, startling changes set in with great rapidity. Around the fibroblasts that are remote from pre-existing normal tissue appear fibrils which at first are delicate and then sturdy in build. The former are demonstrable as fine black lines if the tissue be treated with silver salts for they reduce the salts to silver deposit. They are the familiar reticulin or argyrophilic fibrils, now believed to be the precursors of the coarser collagen fibrils that appear a little later (Dublin, 1946, review). Collagen is found only in immediate contact with cells as a deposit or procollagen, whose outline is determined by the shape of the cell and proximity to other cells. Even far within the haematoma it develops from outlying and often isolated fibroblasts, spreading outwards along the cell fibroglial processes, often lying in close apposition with strands of fibrin. One gets the impression that collagen is solely a matter of fibroblastic activity; certainly it carries with it no suspicion of an origin from fibrin as Baitsell would have us believe. It passes through a stage wherein it is [illegible]

collagen fibrils identical with reticulum except in the property of reducing silver salts. Fusion of fine fibrils or perhaps a kind of fibrillar growth leads to the production of coarse, fully formed collagen with remarkable tensile strength and we now notice a shrinkage and elongation of the fibroblasts which eventually culminates in mature connective tissue. The subsequent history of fibrogenesis is dominated by two mysterious phenomena: (1) orientation to give the peculiar tissue pattern, and (2) reduction in the amount of collagen. About the first pro-

cess I shall have more to say when discussing tendon repair and the second process will be encountered again in the more appropriate place of liver regeneration.

While the fibrous tissue cells are going through their cycles of division and fibrillation, new blood vessels are being modelled from solid endothelial buds with fine terminal prolongations which grow out from pre-existing vessels. The modern investigations of the Clarks (1932, 1933, 1936) and of Florey and his co-workers (Ebert and Florey, 1939; Ebert *et al.*, 1940) with the Clark-Sandison chamber inserted in rabbits' ears and more recently of Algire (1943) using a transparent chamber attached to the mouse's abdominal wall, and of Speidel with his ingenious technique for the tadpole's tail (see 1948) have confirmed the original discoveries of Thoma (Cameron 1952, for discussion) and taught us something about factors therein concerned. New blood vessels grow in the Clark-Sandison chamber at the rate of 0 2–0 6 mm. a day; rate of growth can be modified by temperature and the obstacles to be overcome, especially the consistency of blood clot which may lie in the way of the growing buds. Nuclei increase by mitotic division, blood flows into the lumina that soon form, sprouts join up to form at first loops and later complicated plexuses, while differentiation into venules and arterioles follows in certain places by the alignment of fibrous tissue and muscle cells along endothelium. Where the muscle cells come from is still uncertain. They may grow along the loops from parent vessels or quite likely they represent modified fibroblasts. Equally puzzling are the factors that bring about the orientation of these cells in the vessel walls and attract adventitial cells of reticulo-endothelial origin to the outer limits of the differentiating vessels. And here, too, we encounter forces by which the vasculature may be modelled and remodelled. The lumina of many of the capillaries can be obliterated, the solid cord of endothelium thus produced dividing and the ends retracting to the level of the next patent vessel. In other words, reversibility is a property associated both with the production of fibrils and new blood vessels.

Lymphatics grow in much the same way as the blood capillaries, emerging as solid sprouts from pre-existing lymphatics

(Pullinger and Florey, 1937). Endothelial cells divide and migrate, growing ends anastomose from time to time with one another and even with mature lymphatics, and a lumen is hollowed out along which the lymph soon flows sluggishly. The walls of the lymphatics persist as endothelium without any supporting cells.

Finally, vasomotor nerves can be traced to the vessels in close association with muscle cells and their growth coincides with the assumption of active contraction by the vessels. The earliest contractions were seen by Florey and Ebert in a vessel that had been an endothelial tube 10 days previously, the latest in an arteriole that first contracted 7½ months after its formation.

In healing by first intention the break in the surface epithelium is minute and it is quickly repaired. Should approximation of the edges of a wound not be possible then a much more extensive surface remains to be covered over by epithelium.

The precursory inflammation, often enough complicated by infection, is more severe and leads to the outpouring of exudate with the formation of a surface covering or scab compounded of fibrin, red and white corpuscles and remnants of dead epithelium. Two to three days later, the base of the wound is red and subsequently becomes granular. Each granule represents a newly formed capillary loop surrounded by macrophages, fibroblasts and other inflammatory cells. This young reparative tissue, appropriately labelled granulation tissue, grows into the wound gap, forms fibrils in the manner already described and repairs the lost tissue. Maturation of the fibrous tissue goes on for a long time, and many months may elapse before a tough scar is formed. Even then repair may not be quite perfect for the new tissue shows peculiarities in behaviour on the part of its vessels. Most of us are familiar with the lividity of old scars in cold weather, or their failure to respond to emotion. Surely these idiosyncrasies reflect an unstable vasomotor control. So, too, the embarrassing attribute of scar tissue of grossly contracting, with disfigurement and distortion, indicates imperfection in the control of recovery.

With epithelial repair, however, proliferation and replacement go on rapidly and successfully and the outcome, in most instances, is free from worry. Basal cells from the wound edge

slide across the bare surface and flatten out to cover the gap. Mitotic division also involves the cells some little distance behind the growing edge after the first day or two and this makes available fresh epithelium to bridge the denuded regions. I shall say more about this process when I consider regeneration in the trachea. Then follows more division of cells in their new locations, while differentiation takes place until the original complex cellular covering is restored. At this time or perhaps a little earlier the scab falls off, reparation of the underlying connective tissue ceases and so the wound is repaired, economically and almost completely. Appendages of skin such as hair, sweat and sebaceous glands are not renewed nor are the normal papillae reconstructed. Restoration of the elasticity of the scar is also slow for the making of elastic fibres, whose origin is still uncertain, is one of the last of the differentiations.

Cornea

For an instance of regeneration under simpler conditions we turn to the cornea. Despite the numerous studies on this subject few surpass the well-planned experiments of Senftleben published in 1878. The modern investigations of Beatrice Pullinger and Ida Mann (1943) are equally illuminating and worthy to stand with those of Senftleben. For those of you anxious to pursue the subject further I recommend the authoritative monograph of von Hippel (1937).

Senftleben gently flushed the centre of a rabbit's cornea with a strong solution of zinc chloride so that the epithelial covering is injured without rupture of the outer layers of corneal tissue. Twenty-four hours later destruction of the epithelium and of corneal connective tissue cells has occurred immediately beneath the cauterized area; this is surrounded by a zone of swollen, very distinct corneal corpuscles. Healing then goes on by mitotic division of healthy cells at the periphery of the injury, both the epithelial and corneal fibrous tissue cells participating in division and spreading to cover in the gap left behind as the cauterized cells are discharged. There is no leukocytic migration, no vascular reaction and little oedema. Necrosis is thus followed by repair of the simplest kind.

If more severe injury be produced by deeper cauterization so as to give a break in the surface or by removing a little of the corneal tissue, oedema soon leads to opacity of the underlying cornea, leukocytes migrate from the conjunctiva and corneal spaces to the immediate neighbourhood of the injury and this is followed by repair as in the previous instance. Here again no vascular reactions are apparent and the injury can be repaired without the intervention of new vessel ingrowth from the conjunctiva.

In their beautiful work on healing in the cornea Beatrice Pullinger and Ida Mann (1943) applied minute drops of liquid mustard gas (dichlorodiethylsulphide) to the cornea of rabbits and found that spontaneous avascular healing invariably followed when the centre of the cornea alone was damaged and the corneo-scleral junction was kept uninjured and free from oedema. The initial damage in these lesions entails destruction of epithelium over an area slightly larger than that covered by the application of the liquid. The surface epithelium is quickly wrecked by this highly irritant fluid and corneal corpuscles are destroyed in its pathway as it soaks through the substantia propria into the anterior chamber of the eye. The slit lamp was used to follow changes in the living animals side by side with histological observations at various stages of healing.

Rabbits with a pigmented corneo-scleral junction provide a fascinating demonstration of the sliding or migration of epithelium over the denuded corneal patch, for the pigmented ring, when involved in the lesion, is seen to break, slip in the form of two curving limbs and then scatter over the bare surface. Microscopical examination confirms that this is indeed the result of migration of melanin-containing epithelial cells sliding over the area of damage prior to settling down there to form a surface cover. At the same time the epithelial layer which initiates migration is inundated with inter- and intracellular fluid which can actually be picked up in the slit lamp beam. Pullinger and Mann suggest that this hydropic change may possibly facilitate the slide of epithelial cells.

Meanwhile repair goes on in the substantia propria, preceded by extensive polymorphonuclear cell migration towards the sur-

face of the cornea containing fragments of destroyed corneal corpuscles. This migration is independent of infection and continues for a week or more, the characteristic appearance of the polymorphs giving place to a streamlined arrangement with the cells advancing in Indian file, as if directed or attracted. At the end of a week or ten days, the critical oedema clears away and as the cornea again becomes transparent the normal bright speckled pattern of corneal corpuscles disclosed by the slit lamp beam is replaced by fine white swirling strands resembling floss silk. This represents a second cellular invasion, this time by wandering cells, which starts within 48 hours of injury. Vital staining with Pontamine sky blue allows these wandering cells to be traced from ordinary macrophages concentrated outside the conjunctival vessels. In corneal lesions in which no vascularization takes place the polymorphs and the wandering cells are the only two types that invade the cornea but when blood vessels grow into the cornea they are accompanied by lymphocytes and fibroblasts and other cells of debatable nature and origin called keratoblasts and large hypertrophic cells. Fibrils of collagen are laid down around them and with restoration of the endothelium of Descemet's membrane healing is complete. Opaque white scars, avascular or vascular, sometimes persist for long periods of time according to the severity of the initial destruction.

Tendon

Our next example is regeneration in tendon, a subject that has interested Dr. R. C. Buck of the University of Western Ontario during his recent stay in my department. Buck (1953) followed the changes that occur when the rat's Achilles tendon is divided between its origin and insertion with one stroke of a sharp razor blade and the cut ends are allowed to retract without being sutured. Very quickly the space so formed between the stumps is filled with a fibrinous coagulum in which the fibrin strands lie parallel to the long axis of the defect. During the first 3 days the surrounding connective tissue proliferates to form fibroblasts which infiltrate the fibrin coagulum from without inwards. These young cells soon take on a filamentous appearance and lie with their long axes parallel to the fibrin strands as if they

were apposed to these structures. Within a week or two a band of mature fibrous tissue has filled in the gap of division and is indistinguishable from normal tendon. Where the stumps join this new tissue, fibrous tissue cells are irregularly arranged but are frequently parallel to the cut ends of the tendon. Reticulum and collagen appear about the fourth day and they, too, run in longitudinal parallel lines through the main part of the substitute tendon. They are irregular, transverse or whorled at the stump junctions. After the first fortnight reticulum no longer exists as such but appears to have been converted into collagen. The latter goes on thickening and is very coarse until at least the fourth month after tenotomy.

A metachromatic material, well demonstrated by toluidine blue staining, can be found without much difficulty between the cells and fibres of the paratenon immediately after operation. Buck thinks it represents the normal synovial fluid found in this location. However, in about 4 days a metachromatic ground substance appears between the cells and young fibres of the new tissue and blends with that of the original paratenon. It goes on increasing throughout the new tendon for many months and may be demonstrated even after 12 months.

Fibroblasts appear within the retracted tendon stumps after the 5th day and gradually extend from near the point of section until almost the whole of the tendon, to where it divides into septa at the muscle, is heavily involved. New fibres are formed from these cells and lie between them and the original tendon fibres. They join up with the fibres in the gap of division thereby effecting union between the old stumps and the new tendon. Fibrillar repair seems to go on here for quite a long time since the old stumps remain cellular up to the fourth month.

Blood vessels arising from those of the vascular sheath around the tendon can also be traced into the new tissue about the fifth day and continue to increase over the next 10 days. Simultaneously vascularization is seen in the old stumps and joins the new vessels in the reparative region with those in the muscle tendon junction. In general, vessels grow parallel to the long axis of the tendon but numerous oblique twigs link together the longitudinal tubes. At the site of union with the stumps, vessels

form a plexus and run in all directions. By the end of the third week, a perceptible decrease in the number of vessels is apparent and in later stages only a few lie in the looser parts of the new tendon between condensed bands of collagen. At this time a fibrous sheath forms about the new tendon in which course many vessels. Eventually the regenerated Achilles tendon comes to be composed of a number of large, almost avascular bands separated by generous amounts of loose vascular connective tissue. Buck has shown that the rat is unique in its powers of differentiating cartilage and true bone at the points of junction with the old stumps. Other animals pass through much the same stages of repair as those described by Buck in the rat, but, as far as we can tell from the literature, they show no tendency to develop cartilage and bone.

Many workers have studied tendon regeneration and with few exceptions they have attributed the elongation of the new tendon cells to stretching by the tensional force of muscle pull. Their argument seemed to gain strength from the observation that the earliest fibroblasts are not arranged longitudinally but are disorderly. It was only with resumption of function by the tendon that the cells became elongated. Buck's experience, however, is different. He notes that even the first fibroblasts are somewhat attenuated and that orientation in the plane of the long tendon axis is clear enough. Buck accepts the view of Weiss (1933) who places the main emphasis on initial orientation of fibrin filaments and subsequent spread of cells in close contact with these filaments. Buck maintains that the fibrin in the tendon defect is 'spun out' from blood or exuded plasma under the influence of muscle pull. Along this orientated fibrin scaffolding extend the fibroblasts and they in turn lay down collagen which will therefore be arranged longitudinally. The orientation which the reparative cells and fibrils thus assume is largely determined in the early stages of regeneration and the initial shape of the cells merely reflects their conformity to existing surfaces. Later there is some 'functional adaptation' of cell shape to the mechanical force but even so, Buck reminds us, it is unlikely that the cell is drawn out by tension; rather it is compressed into spaces between the tense fibres.

Finally, Buck confirms the old observation of Lange (1929) that paralysis of the nerve to the muscle whose tendon has been cut favours the production of a tendon-like structure which, however, is much thinner than that formed when the tendon alone is severed. Removal of the bones of the leg at the time of tenotomy leads to a similar result. Under reduced tension the tissue that forms is always as well organized as that which develops under greater tension but there is less of it. Buck suggests that since the new tendon is of smaller diameter, the amount of tension carried by it in proportion to its cross-sectional area is comparable to that of the thicker tendon under greater tension. Whatever may be the explanation of these astonishing variations, clearly it may be concluded that *the amount of reparative tissue which develops is proportional to the functional demand.* I would remind you that quantitative studies of hypertrophy have brought to light an identical rule. I see no reason for separating these two processes for both entail the production of new protoplasm under compulsion.

Trachea

An interesting study of epithelial regeneration has recently been carried out in my department by Dr. D. L. Wilhelm, and as it illustrates very clearly the chief features of restoration of surface epithelium after denudation I shall describe it in some detail.

Wilhelm (1953) has worked out a method for curetting strips of mucous membrane from the trachea of rats using a small aural spoon curette introduced endotracheally without opening that organ. With practice, breaching of the elastic lamina of the basal layer of the mucosa can be avoided and the injury is then truly superficial. Regeneration proceeds in four stages. First of all an aseptic inflammation develops in the tracheal wall beneath the injury, reaching its peak in 24–36 hours, becoming insignificant in 5–6 days and resolving within 10 days. A surface clot of loose fibrin, effused blood and leukocytes forms over the exposed surface, fluid is exuded from the congested vessels of the submucosa and neutrophil leukocytes migrate from these vessels to the site of injury. The fibrin clot generally breaks away at

36–48 hours when regenerating epithelium undermines it, but when the elastic lamina is breached final separation may be delayed to 96 hours.

The second stage of regeneration overlaps the first stage, for it sets in between 2 and 8 hours after injury. Intact epithelial cells at the margins of the wound flatten out, lose their cilia and spread laterally over the raw surface in close association with the elastic lamina. Distances of about 200μ may be covered in 8 hours by these migrating cells. As they clamber along they appear to be preceded by a narrow clear zone in the fibrin they are undermining, as if they were producing a fibrinolytic ferment which clears the way for their progress. Cells also migrate from the necks of curetted submucosal glands. At 24 hours, a *sudden wave of mitosis commences near to the junction with the* spreading cells and extends along the old epithelium for a considerable distance. Both ciliated columnar and basal cells are concerned in mitosis. The wave has passed over by 48 hours and affects the spreading cells very slightly.

The third stage of repair leads to simple stratification of the new epithelium by mitotic activity of the recently formed cells. After 96 hours, the fourth stage is entered upon when the simple stratified epithelium now arranges itself into a layer of low cuboidal cells with an underlying basal zone of flattened cells. By the end of 5 days re-differentiation is well established and within 6 or 7 days low columnar types are often seen. Small intracellular droplets of mucin form in cuboidal cells after 5 or 6 days and constitute well-formed goblet cells after 12–14 days. Ciliated cuboidal or low columnar cells often appear between 10 and 14 days after curettage and are constant after 17–21 days. In 6 weeks re-differentiation is complete, but it may be delayed by severe submucosal inflammation such as results from stubborn infection.

Wilhelm thus lays emphasis on three factors in the epithelial repair of a tracheal defect:

1. Migration of undifferentiated non-ciliated flattened cells derived from the marginal ciliated columnar cells. This would appear to be an example of thigmotaxis, the property whereby the cells, while retaining connexion with one another by their

'hemselves as closely as possible to
937).
flattened migrating cells to the surface so that the number required to bridge the defect is reduced.

3. Mitotic division in the sheets of spreading cells and in the cells at and beyond the edges of the wound.

These factors clearly are similar to those encountered in the healing of skin wounds. They were first stated with precision by Loeb and his colleagues (Loeb, 1897; 1919–20) in their papers on regeneration of squamous epithelium and have been confirmed many times since. You will find comprehensive accounts of all such work in the reviews by Marchand (1901), Arey (1936) and Cameron (1952).

I would like to say, in conclusion, that I regard Wilhelm's method as one of great promise for the study of factors that may influence epithelial repair. In support of this opinion Dr. Wilhelm allows me to mention his recent demonstration that the metaplastic tracheal epithelium, of vitamin A deficiency, repairs a defect in the same fashion as the corresponding normal tissue, a result suggesting that metaplasia does not modify the underlying controlling mechanism of regeneration.

Liver

My last example of regeneration comes from the liver, whose power of recovery from gross injury has been the subject of observation and experimental study for many years. In 1893, Bluck removed up to two-thirds of the rabbit's liver and proved that regeneration is possible, but it is to Ponfick (1889, 1890, 1895) that we owe the proof by measurement that the liver can restore large quantities of lost tissue. In planning his experiment Ponfick first of all found out how much of the organ can be removed with safety. This proved to be no more than four-fifths of the total mass. Three or even four of the major lobes, equivalent to 75 per cent, can be safely taken from both rabbits and dogs. A threefold increase of liver tissue ensued during the next 5–8 weeks which Ponfick attributed to functional stimulation from physiological lack. The newly formed tissue functioned

normally, so Ponfick thought, because it arose from an entirely [illegible] in contrast with the scarred soil in which [illegible] congestion and [illegible] second day onward. The regeneration field is not [illegible] ever, but is at first dotted with pairs of new cells which eventually become so numerous that they form islands penetrating the original tissue in an harmonious sequence. The vascular system regenerates rapidly though the radial arrangement of vessels is often supplanted by a cavernous type. Epithelium of the coarser bile ducts proliferates, producing an unevenness not unlike the ruffles of a collar by the third day. Mitoses appear in the proliferating epithelial cells but no sprouting of ducts is observed. Ponfick placed most emphasis on proliferation of hepatic cells. He noted that the ground plan remained unchanged, since regeneration occurred through interposition. The hypertrophy was built on the framework and foundations of the old lobules and trabeculae. Lobules lost their former regularity and increased in all dimensions. At first growth was slow but after two days it accelerated and this phase extended over 3 or 4 weeks. Ponfick thus was aware of most of the important features of regeneration.

The American embryologist and histologist, F. P. Mall [illegible] that is a model of clarity and a rich source of [illegible] lobules of normal size. [illegible] phery of the lobules around what Mall termed the nodal points, which correspond to the centres of the portal units. It is at these points that growth takes place normally and new lobules are added like leaves to a branch. Quite recently Harkness (1952) has described periportal mitosis during the first day after removal of two-thirds of the rat's liver, an observation that agrees with the views of Mall.

Many fine studies of liver regeneration have been made with human material (Muir, 1908, Herxheimer and Gerlach, 1921) and on animals (Whipple and Sperry, 1909; Fishback, 1929)

which have put the relationship between liver cells and bile ducts in their true perspective. Few observers now doubt that regeneration is mainly the outcome of mitotic division of liver cells, preceded by a latent period of a day or so in which the cells increase only in size but not in number (Brues *et al.*, 1936; Stowell and Lee, 1950; Drochmans, 1950; Harkness, 1951; Yokoyama *et al.*, 1953). The careful analyses of Harkness show that the different components of regenerating liver grow at different rates. Almost every chemical constituent of liver, except glycogen, increases more rapidly than the number of parenchymal cell nuclei. Phospholipid and ribonucleic acid show an early increase which may indicate corresponding increase in the submicroscopic particles of the cells. Harkness has demonstrated, in the rat, delay in the increase of vascular space of the regenerating liver and he suggests that 'growth and division of hepatic cells is a separate process which precedes in time the reorganization of the vascular channel of the liver'. Similar chemical changes are recorded by Stowell and Lee in the regenerating liver of the mouse following poisoning with carbon tetrachloride. They also give information about the behaviour of various enzyme systems such as alkaline phosphatase, and nucleic acids at different stages of early regeneration, but it is difficult at present to fit these facts together into a coherent story. Finally, we are beginning to learn something about the behaviour of organelles during regeneration of liver cells for methods are now being perfected by which the mitochondria and other granular bodies can be isolated from their cells and their number counted. The mitochondrial population of the average normal adult rat cell is approximately 2,500 (Allard *et al.*, 1952) with 118–120 $\times$ 10^9 per gm liver tissue in rats fed a basal diet (Striebach *et al.*, 1953). During regeneration after removal of two-thirds of the rat's liver, it seems that the mitochondria markedly decrease in each cell (Allard *et al.*, 1953). A similar decrease is met with in rat hepatomas. What this may mean we cannot say at present but in view of the importance of mitochondria as cell dynamos we must anticipate a discovery of profound significance to cell physiology, provided, of course, that these very recent observations are confirmed.

There is no gainsaying the extraordinary success of liver restoration after the most diverse insults to the organ. With remarkable promptitude regeneration is set going within a few hours, rapidly reaches a climax which is reflected in the plateau of the curve of regeneration, and in a few weeks, sometimes less, restores the organ to its previous quantitative and qualitative state. Nevertheless, things go wrong with this delicately balanced mechanism from time to time and repair may be inhibited or suppressed. I shall return to this matter a little later on.

Discussion

These five examples must suffice as illustrations of the mechanism of regeneration and I am sure you will agree that there is a remarkable similarity between them. Reduced to its simplest terms repair consists in the replacement of destroyed tissue by new cells—fibroblasts, epithelium of all sorts, endothelium, nervous and related cells—ground substance and fibrils. Individual rules govern the behaviour of each of these elements and so definite are they that we can predict with great accuracy what will happen under most varieties of disturbance. But when we ask what it is that sets going these responses or decides the code of behaviour we quickly find ourselves in difficulties. Even more puzzling is the enigma of control and suspension of repair, *questions intimately bound up with many fundamentals of* pathology, not the least being neoplasia.

For many years now, biologists have intuitively accepted the idea that damaged or dead tissue furnishes something that stimulates living tissue to replace that which was lost. It is to this belief that I wish to devote the remainder of my lecture. Such *growth-promoting substances*, regeneration hormones, wound hormones, trephones, have been postulated by Marchand (1901), Bier (1917), Haberlandt (1921), Carrel (1924) and many others, often, I fear, from insufficient evidence. The most careful study in recent years of the healing of experimental wounds has been made by Young *et al.* (1941) and this work merits description.

By means of special trephines, 12–16 mm. diameter, standard wounds were made in the depilated, sterilized skin of rabbits by

removing discs of skin and subcutaneous tissue, clipping away the areolar tissue from the underlying aponeurosis and trimming the margins of the wound to eliminate shelving.

Bleeding was controlled and the wounds were allowed to heal under a scab, in most cases without any dressing. Sepsis seldom occurred. Tracings of the wounds were made and their areas measured from day to day on cellophane until every granulating point had been covered with epithelium and healing was complete. Large numbers of animals were used and the results were analysed by the late Dr. Matthew Young, a statistician of great experience. The main aim was to find out whether the rate

... by the presence of another

... ngle wounds, others paired,

triple and quadruple ... econdary wounds, the latter or test wounds being made at an interval of 10 or 12 days after the primary wounds. Due allowance was made for the well-established fact that large wounds heal at a greater rate than smaller wounds and that, as a rule, the rate of closure is directly proportional to the area of the wound.

The results were definite enough. Secondary wounds closed at a greater rate than primary wounds and in seven out of the ten size-groups—the heterogeneous series of primary and secondary wounds had to be classified as groups—the difference in rate was statistically significant. Hence it was concluded that some kind of accessory accelerating factor must operate in the healing of secondary wounds which is lacking in relation to primary wounds. With admirable caution Young *et al.* discuss the nature of this factor but they cannot decide whether it is truly a growth-promoting substance liberated by the primary wounds, a by-product of immunity, or some other principle. They feel certain, however, that this accessory factor is less potent than the size factor.

So far as I know, this is the first properly conducted quantitative study of growth-promoting factors in wound healing.

... to it is seen in the experiments of

...

wounds increased ...

described analogous results with liver regeneration and Newman *et al.* (1949) also claim that regeneration is more rapid in rats fed on powdered whole liver or freshly cooked whole pig's liver, than with other, non-specific proteins. Teir *et al.* (1951) recently have shown that the intraperitoneal injection into rats, every other day, of a skin extract accelerates healing of a standard skin defect as compared with untreated controls. Teir and Ravanti (1953) also show that suspension of fresh liver injected intraperitoneally into rats stimulates mitosis in the liver. These observations are suggestive and full of interest, but many more of the same sort of experiments are needed before we can draw widesweeping conclusions.

Brief mention may be made of a number of other factors which influence healing before we turn to the discussion of cellular behaviour. *Age* plays an important part in repair of human and animal wounds (du Nouy, 1916; Howes and Harvey, 1932), for wounds heal with greater rapidity in the young than in adults. The following summary of work by Bourlière (1940) shows this very neatly (Table 1). The standard skin wounds in their experiments were made on the posterior body wall over the midline of female rats. Diet and temperature were carefully controlled throughout the experiments.

TABLE 1

Group	Age in months	Number of rats	Mean weight, grams	Mean initial surface of wound, cm.²	Mean time of healing in days
1	3–7	23	134 ± 26	2·48 ± 0 33	17·8 ± 1 6
2	10–14	19	163 ± 30	2 55 ± 0 30	21 0 ± 3·1
3	18–30	22	189 ± 23	2·70 ± 0 45	21 6 ± 2 6

Healing was significantly more rapid in young animals than in adults, but old animals did not differ from adults in their capacity to heal A similar law holds for regeneration in the liver after its partial removal. I found this to be the case with rabbits in experiments performed 20 years ago, but never published. More recently, Norris *et al.* (1942) removed 65 per cent of the liver from rats and showed that young animals are better at

removing discs of skin and subcutaneous tissue, clipping away the areolar tissue from the underlying aponeurosis and trimming the margins of the wound to eliminate shelving.

Bleeding was controlled and the wounds were allowed to heal under a scab, in most cases without any dressing. Sepsis seldom occurred. Tracings of the wounds were made and their areas measured from day to day on cellophane until every granulating point had been covered with epithelium and healing was complete. Large numbers of animals were used and the results were analysed by the late Dr. Matthew Young, a statistician of great experience. The main aim was to find out whether the rate of healing of a wound is influenced by the presence of another wound. Some of the rabbits bore single wounds, others paired, triple and quadruple primary and secondary wounds, the latter or test wounds being made at an interval of 10 or 12 days after the primary wounds. Due allowance was made for the well-established fact that large wounds heal at a greater rate than smaller wounds and that, as a rule, the rate of closure is directly proportional to the area of the wound.

The results were definite enough. Secondary wounds closed at a greater rate than primary wounds and in seven out of the ten size-groups—the heterogeneous series of primary and secondary wounds had to be classified as groups—the difference in rate was statistically significant. Hence it was concluded that some kind of accessory accelerating factor must operate in the healing of secondary wounds which is lacking in relation to primary wounds. With admirable caution Young *et al.* discuss the nature of this factor but they cannot decide whether it is truly a growth-promoting substance liberated by the primary wounds, a by-product of immunity, or some other principle. They feel certain, however, that this accessory factor is less potent than the size factor.

So far as I know, this is the first properly conducted quantitative study of growth-promoting factors in wound healing. Perhaps the nearest approach to it is seen in the experiments of

The role of *vitamin A* for epithelial regeneration and differentiation is well established in skin and mucous membranes, but little is known about its role elsewhere. However, Wilhelm has recently shown that deficiency of vitamin A does not retard repair of the tracheal epithelium, On the other hand, Sabella *et al.* (1951) state that the local application of vitamin A to the skin of rats induces epidermal thickening to twice normal. The stratum granulosum increases in size but keratin is formed more slowly.

The necessity for a proper *blood supply* during regeneration needs no emphasis. Wounds that are poorly nourished heal slowly and imperfectly, are often oedematous and of lowered tensile strength. Moreover they easily become the site of infection which not only throws back healing but adds its own special problems. For quantitative investigations of the part [illegible] Magath, 1922; Stephenson, 1932; Higgins *et al.*, 1932; Mann, 1944; Grindley and Bollman, 1952).

I pass on rapidly over the problematical factors of nerve supply, temperature and ultra-violet light, despite their interesting implications, and conclude by referring to *mechanical influences*. The studies on wounds and tendon that I have described to you show how necessary it is to remove obstacles in the way of the regenerating tissue. Surely the intense phagocytic and enzymatic activity accompanying repair is designed to cope with this task. Even more remarkable instances are to be seen in the regeneration of muscle and nerve fibres, where the success of the process may be completely interrupted by imperfectly prepared pathways along which the new muscle cells or nerve fibres should grow. Indeed, regeneration may be held up completely, or diverted to unwonted and harmful directions. The need for suitable surfaces along which epithelium can grow is shown very beautifully in Wilhelm's experiments on the trachea. I suspect that in the liver, too, architectural factors are of vital importance. The arrangement of liver cells is so complex and the ease with which the lobular architecture may be distorted so

restoring their liver mass and forming new liver cells than adults. Old rats equalled adults in ability to restore the liver mass but lagged behind somewhat in cell production. Bucher and Glinos (1950) have demonstrated identical rules in a most careful study of regeneration in rats. We have no idea what is the real nature of this age difference; it seems to be part of a general growth phenomenon in which far-reaching influences are involved.

Another factor of great importance in regeneration is *food supply*. A certain amount of latitude seems to be allowed for some animals, e.g. rats and certain invertebrates, which can still regenerate normally when starved for short periods. With prolonged starvation, however, repair is inhibited and may even stop. I have discussed the rather scanty and conflicting information about this topic elsewhere (Cameron, 1952) and will content myself by mentioning a careful study by Morris *et al.* (1944–5). These workers have shown that a diet with a high protein content increases the rate of epithelialization of a skin wound in rats. This increase is attained even when the animals have been fed a low protein diet provided they are placed on the high protein diet at the time of wounding. Obviously the fed protein can be made use of at once in repair. The addition of the amino acids valine, tryptophane and lysine to the low protein diet does not appreciably affect retarded epithelialization. Attention may also be drawn to the claim of Clark (1919) that diet moderates only the latent period of wound healing in dogs. A high protein diet seems to eliminate the latent period completely, a high fat diet prolongs it. This seems to be a profitable field of investigation.

The importance of *vitamin C* has been stressed in describing the formation of collagen and ample confirmation of this relationship has come from many careful studies of repair (Cameron, 1952).

The rate of regeneration of wounds in corneal epithelium or of epithelium of gum, thigh or ear pinna of the guinea-pig is not retarded by lack of vitamin C But the slow formation of collagen in the healing wound when ascorbic acid is deficient delays epithelialization and retards wound healing as a whole (Galloway *et al.*, 1948–9).

results, this time based on estimations of wet weight and number of liver cells, have been recorded by Wennekcr and Sussman (1951). Whether this is a specific phenomenon or not we do not know. Experiments carried out by Marshak and Walker (1945) show that mitosis in regenerating liver is increased if the animals are given intravenous injection of small amounts of chromatin which rather suggests a non-specific action.

The old work of Wells (1906) which was confirmed in unpublished experiments by Oakley and myself points to a product of autolysing tissue as a growth-stimulator of connective tissue cells and vascular endothelium. If a piece of liver be implanted in the peritoneal cavity or omentum of the same animal it autolyses and induces around it a pronounced reparative reaction in which fibrous tissue and blood vessels develop within a few weeks. By coagulating the liver fragment before implantation, autolysis is largely prevented, and now the reparative reaction is delayed often for many weeks or months. Clearly a product of autolysis is responsible for the new formation of fibroblasts and endothelium in the first instance, which is followed by their differentiation into mature scar tissue. The experiment works with all sorts of tissues. I would suggest, therefore, that the reparative mechanism consists in the production, by enzyme action, of one or more chemical compounds which promote the division of parenchymal as well as stromal cells and thus bring about balanced replacement of lost tissue. I would like to emphasize that this suggestion is well-nigh pure hypothesis but I believe it to be a useful suggestion for further work. A similar suggestion has been made about the proliferation of Schwann cells preliminary to nerve regeneration by Abercrombie and Johnson (1942). And this also may be said in its support, that we are at last becoming familiar with a variety of agents that influence the mitotic activity of cells *in vivo* as well as *in vitro* (Bullough, 1949; 1950). Some stimulate division, others inhibit. Sex hormones—oestrone, testosterone—and thyroid extract are examples of the first group, adrenal cortical hormones, especially cortisone, are inhibitors. And that brings me to a consideration of hormones on other activities of cells concerned in regeneration.

apparent that I can well believe that newly formed cells may find great difficulty in retreading the old and proper pathways to restore the lost lobules. The liver seems to have got over this difficulty by forming new cells in all sorts of places, particularly at nodal points where growth normally goes on and where there is ample room for the laying down of new lobules to replace those lost.

This survey of the main factors influencing repair of tissues now encourages me to attempt the *final analysis of the regenerative responses* of cells and their derivatives. Once again I must warn you that information is scrappy and often woefully deficient but surely that should not deter us from making the attempt to see ahead a little further.

I have already mentioned the demonstration by McJunkin and Breuhaus that liver growth may be stimulated by feeding rats with fresh liver. These workers based their conclusions on the finding of many more mitotic figures within liver cells of experimental animals than of control animals and this would seem to suggest that such cells were in fact proliferating. More satisfactory grounds for the belief that something is produced after injury of liver cells that stimulates healthy cells to divide comes from the ingenious experiments of Aub and his colleagues (Bucher *et al.*, 1950, 1951). One partner of each of a number of pairs of parabiotic Wistar rats was partially hepatectomized and the livers of the non-hepatectomized partners scrutinized after 2 or 3 days for evidence of growth by counting mitoses. Sure enough a higher mitotic rate was found in the liver of the non-hepatectomized animals than in the liver fragments removed from the other partner at the time of excision. The conclusion seems justified that mitotic activity during regeneration, at any rate in the liver, is initiated by alteration in the chemical composition of the blood and that probably a stimulating agent is produced because of injury. If this be the case, the agent must surely be very powerful for partial hepatectomy when carefully performed leaves behind relatively little dead or damaged tissue. Of course such a property is an essential requirement for any hypothetical growth-stimulator if it is to be effective after the extremes of injury that are encountered in animals. Similar

in various organs as the result of a pathological process can also be removed in part or completely. Some years ago Professor W. A. E. Karunaratne and I (Cameron and Karunaratne, 1936) showed that reticulin and collagen scars can disappear from the livers of rats in which we had produced cirrhosis by the prolonged administration of carbon tetrachloride. Orr (1940) and Morrione (1947) have confirmed our results, the latter using methods for the estimation of collagen in organs. Here is a summary of his results.

TABLE 2. Reversal of Carbon Tetrachloride Cirrhosis in Rats, as shown by Estimation of Liver Collagen (after Morrione)

No. of rats	Rat weight (g.)	Liver weight (g.)	Hepatic collagen weight (per cent)	Total hepatic collagen (mg.)
Cirrhosis (inhalation of carbon tetrachloride 3 or 4 times weekly for 3½ months)				
25	181 6	8 6	1 01 ± 0 14	86·4
Reversal (carbon tetrachloride discontinued for 1 month)				
10	234·5	10 4	0 69 ± 0·17	73 7
Reversal (carbon tetrachloride discontinued for 2 months)				
12	223 8	9·9	0 61 ± 0 15	59 6
Reversal (carbon tetrachloride discontinued for 3 months)				
9	270 7	12 4	0 51 ± 0·01	63·0

Again we are in the presence of a mystery which may eventually be resolved in terms of some of the suggestive facts I have already put before you. When we know the answer we shall be on the threshold of a profound advance in biological theory and, I fancy also, of a move towards the cure of many chronic diseases.

REFERENCES

ABERCROMBIE, M and JOHNSON, M L. (1942) *J exp. Biol* 19, 226.
A[illegible]
[illegible]
Cancer Res. 12, 407

Space does not permit of a full discussion of this important topic, nor, indeed, are we all too sure of our facts. This much may be claimed at present, that fibrous tissue production is stimulated by desoxycorticosterone, thyroxine and pituitary growth hormone and inhibited by cortisone, testosterone and oestradiol (references in Cameron and Spector, 1954). Cortisone depresses fibroblastic and angioblastic activity and reduces the amount of ground substance by inhibiting, amongst other things, the synthesis of chondroitin sulphate. Desoxycorticosterone stimulates fibroblastic function and encourages collagen formation. In this way it resembles ascorbic acid. Even the ground substance or cementing material of tissues seems to come under the control of hormones, for its water-binding properties can be increased by testosterone and decreased by some female sex hormones. I shall not attempt to discuss the position of the hyaluronidases for the matter appears to be in utter confusion. Finally we have a little evidence of similar processes at work with epithelial cells, especially in the relationship on the one hand between thyroid, pituitary and perhaps adrenal cortical hormones, and the liver on the other hand. But contradictions are all too numerous and the separation of general growth effects from specific regenerative activity is so uncertain that we shall do well to await further information before adopting a set hypothesis.

In conclusion I would remind you that the business of regeneration, so far as we understand it in animals, is not a fixed, unalterable mode of procedure, adjusted to a rigid pattern with laws of qualitative and quantitative behaviour. For a long time we have been familiar with the workings of two antithetical biological processes, atrophy and hypertrophy, whereby tissue bulk can be reduced or increased according to the requirements of the animal. Experience is teaching us that either may be reversed provided it has not gone too far and indeed one may be converted into the other, almost at will. And this is so with regeneration. A wound in healing often produces a great deal more reparative tissue than it apparently requires, for much of the connective tissue and many of the blood vessels may disappear in time. There is reason to believe that scars laid down

MANN, F. C. and MAGATH, T. B. (1922) *Amer. J. Physiol.* **59**, 485.

MORRIONE, T. G. (1947). *J. exp Med.* **85**, 217.

MORRIS, H. P., DUBNIK, C. S. and DUNN, T. B. (1944–5). *J. Nat. Cancer Inst.* **5**, 271.

MUIR, R. (1908). *J. Path. Bact.* **12**, 287.

NEWMAN, E., GROSSMAN, M I. and IVY, A. C. (1949). *Amer. J. Physiol.* **157**, 221.

NORRIS, J. L, BLANCHARD, J. and POVOLNY, C. (1942). *Arch Path* **34**, 208.

ORR, J. W. (1940). *J. Path Bact* **50**, 393.

PONFICK, E. (1895). *Arch. path. Anat.* **138**, suppl. p. 81.

PONFICK, E. (1889) *Jahrb. Schles. Gesell vaterl. Kult.* **67**, 75.

PONFICK, E. (1889). *Arch. path. Anat.* **118**, 299; **119**, 193

PONFICK, E. (1890). *Verhandl. dtsch Gesell. Chirurg.* **19**, 28.

P ... 15, 157.

SABELLA *et al* (1951). Quoted by Cameron and Spector

... *er Res.* **13**, 279.

... 51.

TEIR, H. and RAVANTI, K. (1953). *Exper. cell Res.* **5**, 500.

VON HIPPEL, E. (1937) Vol. XI, pt. 1, p 261 in Henke-Lubarsch's *Handbuch*

WIGGLESWORTH, V. B. (1937) *J. exp. Biol* **14**, 364.

WILHELM, D. L (1953). *J. Path. Bact.* **65**, 543.

WOLBACH, S. B (1933). *Amer. J. Path* **9**, 689.

YOKOYAMA, H O., WILSON, M. E, TSUBOI, K. K and STOWELL, R. E. (1953). *Cancer Res* **13**, 80.

YOUNG, M., YOUNG, J S. and FISHER, J A. (1941) *J. Path. Bact* **52**, 225

... 16 ...

... 658

BUCHER, N. L. R. and GLINOS, A. D. (1950). *Cancer Res.* **10**, 324.
BUCHER, N. L. R., SCOTT, J. F. and AUB, J. C. (1950). *Cancer Res.* **10**, 207.
BUCHER, N. L. R., SCOTT, J. F. and AUB, J. C. (1951). *Ibid.* **11**, 457.
BUCK, R. C. (1953). *J. Path. Bact.* **66**, 1.
BULLOUGH, ...
BULLOUGH, ...
BULLOUGH, ...
... p. 443.
... **42**, 1.
... *Trends in*

DUBLIN, W. B. (1946). *Arch. Path.* **41**, 299
DU NOÜY, P. L. (1916) *J exp. Med.* **24**, 461.
EBERT, R. H. and FLOREY, H W. (1939). *Brit J exp. Path* **20**, 342.
EBERT, R H, SANDERS, A. G and FLOREY, H. W. (1940). *Brit. J. exp. Path.* **21**, 212.
FISHBACK, F. C (1929) *Arch Path* **7**, 955.
GALLOWAY, N. M., GARRY, R C and HITCHIN, A. D (1948–9). *Brit. J Nutrition*, **2** 226.

HARKNESS, R. D. (1952). *J Physiol* **117**, 267
HERXHEIMER, G. and GERLACH, W. (1921). *Beitr path Anat.* **68**, 93.
HIGGINS, G. M., MANN, F. C. and PRIESTLEY, J T. (1932). *Arch Path* **14**, 491.
... *Med* **55**, 577.

... *J med. Res* **41**, 247.
... 1931). *Arch Path.* **12**, 900.
... *Arch. Path.* **12**, 794.
... 7.
... **11**, 65

intercellular substances, so that exploration of their possible biological functions has become a practical proposition and the hypothesis that their derangement may give rise to a variety of diseases can now be tested. This field of study is still very new so that no authoritative answers can be given, but it may be profitable to discuss some of the recent work and to see whether it provides a useful approach to the study of the many diseases in which connective tissue changes figure prominently.

THE UNITY OF THE SYSTEM

Before discussing the various components of connective tissue, it may be as well to consider the supporting system as a whole lest we lose sight of the wood through preoccupation with the trees. Our supporting framework would seem to be made of diverse structures such as bone, tendon, cartilage; fascia and synovium; the dermis of the skin, the sclera and cornea in the eye, heart valves and the walls of blood vessels and the reticulum of the more cellular organs; but there is an essential unity between all these structures since they are all connective tissues, derived from the same mesenchymal cells which persist throughout life in the granulation tissue that follows injury of any kind. These structures are also composed of similar mucoid and fibrous intercellular material; the apparent differences between one structure and another being largely the result of differing proportions and modes of organization of these materials. Thus information derived from a study of the cornea may have a bearing on the problems of articular cartilage and vice versa, and an understanding of disorders of ligament or dermis may throw light on disorders of the heart valves Certainly the clinician who interests himself in diseases involving connective tissue cannot help being struck by the frequency with which these diseases involve many structures such as the skin and eye as well as the cardiovascular system and the bones, joints and tendons, and I believe that if we develop the habit of looking at the supporting system as a whole it may lead to a useful synthesis of much scattered and apparently unrelated information.

V

The Supporting System and its Disorders

J. H. KELLGREN

Historical

After the introduction of the cell doctrine by Schwann in 1839 it became customary to think of the human body as a mass of living cells. Though there is much truth in this concept, it is only half the truth since more than half the body is composed of extracellular material. Part of this material between cells is fluid and is concerned with internal transport and homeostasis, but much of it consists of fibres, mucoid and cement, forming the supporting framework which gives our bodies cohesion and form, and a system of rigid levers with which mechanical forces can be applied and resisted.

Although collagen and elastic fibres were characterized histologically and to some extent chemically during the nineteenth century (Rollett, 1871), the personality of Virchow and his dictum 'Omnis cellula et cellula' remained dominant, and the extracellular framework was regarded as so much inert stuffing that could have little bearing upon the problems of health and disease. It was not until 1933, when Klinge showed that alterations in the intercellular substances were the most striking feature of certain rheumatic diseases, that interest in these substances was reawakened, and since then the concept of collagen diseases (Klemperer, 1950), or connective tissue diseases (Kellgren, 1952) has received increasing attention and it has even been suggested that the state of the extracellular ground substance may control cellular activity (Gersh, 1950).

Recent developments in biophysics and biochemistry have provided new methods for studying the somewhat complex

of the ground substance as in cartilage. It is mainly polysaccharide but with a small protein fraction. The cement is the material that binds the fibrils into fibres and forms non-fibrillar structures like basement membrane; though it has not been fully characterized, it is probably mucoprotein in nature, that is mainly protein but with a large carbohydrate component.

The fibre system provides strength and rigidity, particularly when impregnated with mineral salts as in bone, while the mucoid provides lubrication and a certain elasticity and resilience to the fibre systems. Both mucoid and cement also appear to be important stabilizers of the fibre system and to be concerned with tissue permeability and possibly ion exchange and other functions, and a component of basement membrane is [illegible]

in which of course the supporting system is mesodermal in origin, but it is interesting to note that in arthropods (Richards, [illegible]

matrix which may become more fixed and rigid by a process of quinone tanning or by impregnation with mineral salts, although calcium carbonate in the form of calcite is then the main component instead of calcium phosphate in the form of apatite as in vertebrates.

Types of Fibres

Returning to man, we find three main types of connective tissue fibres, collagenous, reticular and elastic. Collagen is the main component of white fibrous tissue; under the light microscope this appears as wavy unbranched longitudinally striated fibres of differing thickness, which are birefringent and therefore beautifully displayed by polarized light. They stain characteristically red with Van Gieson's picro-fuchsin. The electron microscope reveals unbranched and cross-banded fibrils of varying thickness and with varying spacings between the cross-bands

THE CELLS

Although the connective tissue cells presumably play an important role in the formation and maintenance of the supporting system, surprisingly little is known about their activities. In bone the dominant part played by osteoblasts and osteoclasts seems well established but no such remodelling cells have been described in other tissues. It is however known that collagen fibres can be produced by fibroblasts in tissue culture (Doljanski and Roulet, 1933; Porter and Vanamee, 1949) and that the matrix surrounding cartilage cells will disappear if the tissue is grown in a medium containing excess of vitamin A (Fell and Mellanby, 1952), and that fibre formation is defective in the absence of ascorbic acid (Wolbach, 1933); but virtually nothing is known about the enzyme systems by which the connective tissue cells affect the surrounding matrix. On the other hand the supporting system as a whole is very responsive to mechanical forces. Thus shearing strains will produce a synovial cavity, compression stimulates bone formation (Charnley, 1953), and distraction favours the formation of fibrous tissue; while the absence of all mechanical stimuli leads to rapid atrophy. In articular cartilage intermittent compression appears to condition the degree of hydration (Ekholm and Norbäck, 1951) and the chondroitin content (Matthews, 1952) as well as the rate of flow of colloid particles from the epiphyseal capillaries through the cartilage into the joint cavity (Inglemark, 1950). But to what extent these changes are produced by the direct action of mechanical forces upon the intercellular substances or are mediated through cellular mechanisms remains to be determined.

THE FRAMEWORK

The intercellular portion of the supporting system has three main components fibres, mucoid and cement. The fibrous material is composed of rather stable, insoluble proteins with a very small carbohydrate fraction and they are arranged in different ways: as parallel bundles of fibres as in tendon, or as a criss-cross lattice as in cartilage and skin, or as lamellae as in bone and arterial walls. The mucoid is either a viscous lubricating solution as in synovial fluid or it may be the main constituent

of the ground substance as in cartilage. It is mainly polysaccharide but with a small protein fraction. The cement is the material that binds the fibrils into fibres and forms non-fibrillar structures like basement membrane; though it has not been fully characterized, it is probably mucoprotein in nature, that is mainly protein but with a large carbohydrate component.

The fibre system provides strength and rigidity, particularly when impregnated with mineral salts as in bone, while the mucoid provides lubrication and a certain elasticity and resilience to the fibre systems. Both mucoid and cement also appear to be important stabilizers of the fibre system and to be concerned with tissue permeability and possibly ion exchange and other functions, and a component of basement membrane is [illegible]

in which of course the supporting system is mesodermal in origin, but it is interesting to note that in arthropods (Richards, 1951), in which the supporting system is cuticular and therefore ectodermal in origin, the insoluble fibre system is formed by the polysaccharide chitin which is set in a more soluble protein matrix which may become more fixed and rigid by a process of quinone tanning or by impregnation with mineral salts, although calcium carbonate in the form of calcite is then the main component instead of calcium phosphate in the form of apatite as in vertebrates.

Types of Fibres

Returning to man, we find three main types of connective tissue fibres, collagenous, reticular and elastic. Collagen is the main component of white fibrous tissue; under the light micro- [illegible] gitudinally striated

fringent and there-

They stain charac-

chsin. The electron

microscope reveals unbranched and cross-banded fibrils of varying thickness and with varying spacings between the cross-bands

up to 650 Å, depending upon the anatomical site and the maturity of the tissue examined (Randall *et al.*, 1952). X-ray diffraction at high angles gives a characteristic pattern with a meridional arc at 2·86 Å and equatorial spots at 10 to 15 Å, the latter being influenced by the degree of hydration. At low angles longer spacings corresponding to the cross-bands revealed by the electron microscope are obtained (Bear, 1952). On boiling, collagen goes into solution forming gelatin. Thus there are many possible methods of defining 'collagen' and serious semantic difficulties arise from the fact that these are not all co-extensive.

Reticulin forms the reticulum of the cellular organs. Under the light microscope it appears as fine branched anastomosing fibres that are not birefringent, and which characteristically stain black with silver. The electron microscope reveals cross-banded fibrils similar to those of collagen but set in amorphous sheets or membranes. X-ray diffraction gives the pattern of disoriented collagen. On boiling, the fibrillar portion of reticulin goes into solution but does not form a gel, while the amorphous portion remains as a precipitate (Kramer and Little, 1953). Rather similar argyrophil fibres may be seen throughout young growing fibrous tissue, and in regions of fibrous tissue breakdown, but this material seems to be very closely related to collagen and may not be identical with the reticulum of cellular organs.

Elastin is the main component of the yellow elastic tissue found in certain spinal ligaments and in arterial walls though some elastin is scattered throughout most of the connective tissues.

Under the light microscope it appears as smooth branching refractile fibres as in the ligaments or connective tissues or as fenestrated membranes as in arterial walls. It stains darkly with orcein or resorcin-fuchsin. The electron microscope reveals smooth branching fibrils and elastin appears to be completely unaffected by boiling.

Collagen, elastin and reticulin may also be defined in terms of their amino-acid content (Bowes and Kenten, 1949; Lansing, 1951). All three contain large amounts of the non-polar amino acids glycine and alanine and substantial amounts of proline.

Collagen contains very few aromatic residues and is the only protein with a substantial content of hydroxyproline so that the collagen content of tissues may be estimated in terms of hydroxyproline. But reticulin also has a high content of hydroxyproline (Kramer and Little, 1953) and although the amino-acid composition of reticulin has been less fully worked out it appears to be very similar to that of collagen. Furthermore, reticulin resembles collagen in that it is digested by collagenase and pepsin while being resistant to trypsin and also elastase which selectively digests elastin.

Thus collagen and reticulin are similar in many respects whereas elastin would appear to be something quite distinct.

Mucoid

The mucoid is not readily seen in ordinary histological preparations but it is revealed by special staining techniques, showing metachromasia with toluidine blue and staining deeply with the Hale and periodic-acid Schiff techniques. It is largely composed of polysaccharides such as hyaluronic acid and chondroitin sulphate. The former seems to be mainly a lubricant and is readily obtained from synovial fluid while the latter is mainly obtained from the matrix of cartilage, and appears to be more of a stabilizer of the fibrous lattice. There may also be other polysaccharides, notably in subcutaneous tissue (Consden *et al.*, 1953) and in cornea (Woodin, 1952), which have not yet been fully characterized.

The polysaccharides are generally described as long-chain polymers of disaccharides composed of glucuronic acid and various hexosamines (Meyer, 1950), and tissue polysaccharide is often estimated in terms of hexosamine content. The capacity of polysaccharides to form viscous lubricating fluids, or a stabilizing tissue matrix, depends upon their degree of polymerization rather than the total content which is important. Though certain testicular and bacterial extracts depolymerize polysaccharides, and these enzymes have been used to facilitate the spread of fluid through connective tissues, such enzymes have not yet been extracted from connective tissue cells.

up to 650 Å, depending upon the anatomical site and the maturity of the tissue examined (Randall *et al.*, 1952). X-ray diffraction at high angles gives a characteristic pattern with a meridional arc at 2·86 Å and equatorial spots at 10 to 15 Å, the latter being influenced by the degree of hydration. At low angles longer spacings corresponding to the cross-bands revealed by the electron microscope are obtained (Bear, 1952). On boiling, collagen goes into solution forming gelatin. Thus there are many possible methods of defining 'collagen' and serious semantic difficulties arise from the fact that these are not all co-extensive.

Reticulin forms the reticulum of the cellular organs. Under the light microscope it appears as fine branched anastomosing fibres that are not birefringent, and which characteristically stain black with silver. The electron microscope reveals cross-banded fibrils similar to those of collagen but set in amorphous sheets or membranes. X-ray diffraction gives the pattern of dis-oriented collagen. On boiling, the fibrillar portion of reticulin goes into solution but does not form a gel, while the amorphous portion remains as a precipitate (Kramer and Little, 1953). Rather similar argyrophil fibres may be seen throughout young growing fibrous tissue, and in regions of fibrous tissue break-down, but this material seems to be very closely related to collagen and may not be identical with the reticulum of cellular organs.

Elastin is the main component of the yellow elastic tissue found in certain spinal ligaments and in arterial walls though some elastin is scattered throughout most of the connective tissues.

Under the light microscope it appears as smooth branching refractile fibres as in the ligaments or connective tissues or as fenestrated membranes as in arterial walls. It stains darkly with orcein or resorcin-fuchsin. The electron microscope reveals smooth branching fibrils and elastin appears to be completely unaffected by boiling.

Collagen, elastin and reticulin may also be defined in terms of their amino-acid content (Bowes and Kenten, 1949; Lansing, 1951). All three contain large amounts of the non-polar amino acids glycine and alanine and substantial amounts of proline.

of mature tissues is really a complex of a more soluble type of collagen plus a small fraction of chondroitin or other polysaccharide and a tyrosine-containing protein or polypeptide, the latter fractions being possibly combined into a mucoprotein cement substance.

It has been suggested that elastic fibres also have a dual composition, being formed of linear aggregates of a corpuscular protein held together by a cement substance which contains both polysaccharide and sulphuric acid (Hall *et al.*, 1952).

Fibrinoid

Fibrinoid connective tissue was first described by Neumann (1880), and when Klinge demonstrated its widespread occurrence in many rheumatic diseases the nature of this material became of general interest. In rheumatoid disease fibrinoid occurs in minute spots scattered sparsely throughout the affected tissues, but on the elbows and elsewhere sizeable aggregates may form which, with their surrounding zones of fibrous and inflammatory tissue, constitute the classical nodules.

Fibrinoid areas can often be seen with the naked eye as yellowish opaque spots in the semi-translucent connective tissue but they can more certainly be recognized by histological investigation, when they are distinguished by the following characteristics.

In a fibrinoid area connective tissue cells are few or altogether absent, and the fibre pattern of normal connective tissue is replaced by a disorderly mess of indistinctly fibrillar or altogether amorphous material which stains more intensely with eosin than the surrounding connective tissue. In a structure like tendon a whole tendon bundle may be replaced by such granular amorphous material.

In sections treated by Lillie's reticulin method and Van Gieson's stain the edge of a fibrinoid area shows normal red collagen fibres passing into fine black argyrophil reticulin fibres. There is also an increase of amorphous yellow-staining material which may altogether replace the fibre pattern Those parts that contain silver-staining fibres also stain deeply with the periodic-acid Schiff technique, but the totally amorphous material stains less strongly by this method.

Cement

The supporting framework also contains a small quantity of material which the older histologists described as cement and basement membrane. The exact nature of this material is not known though from its staining properties it may be expected to contain a substantial carbohydrate fraction; though it may be related to the amorphous component of reticulin, in chemical fractionations of connective tissue a residue is obtained which is mainly protein in nature, and contains substantial quantities of tyrosine, little if any hydroxyproline and much hexosamine (Consden *et al.*, 1953). No accurate histologico-chemical correlation has yet been completed but there does appear to be a small mucoprotein or glycoprotein component in connective tissue which may be biologically interesting; especially since this fraction may be increased in rheumatic disease.

The Tissue Complex

Although these various components can be extracted from connective tissues, the tissue itself is really a complex the properties of which are determined by the interaction of these various components. Thus, if cartilage is treated by $CaCl_2$ which extracts the chondroitin matrix (Partridge, 1948) some collagen also goes into solution. If the small chondroitin component of tendon is removed by treatment with hyaluronidase the tendon collagen becomes more soluble and its shrinkage temperature, which is a useful index of fibre stability, becomes much reduced; and *in vitro* chondroitin-precipitated collagen is more stable than salt-precipitated collagen (Jackson, 1953). The treatment of hide by alkali also renders the hide collagen more soluble and this treatment removes some material with a high tyrosine content as well as polysaccharide. Furthermore, a more soluble form of collagen can be extracted from the tissues of young growing animals by citrate buffers and this citrate-extracted collagen has a lower hexosamine content, a lower content of tyrosine, histidine, lysine and proline and a higher content of hydroxyproline, alanine and serine than ordinary collagen fibres (Bowes *et al.*, 1953).

These various findings suggest that the stable collagen fibre

Metabolism

[illegible]

in rats are of great interest. These studies of collagen metabolism using ^{14}C-labelled glycine have shown that the main mass of collagen in adult animals is very inert but that in young actively growing animals there is a small metabolic turnover (Neuberger, Perrone and Slack, 1951; Neuberger and Slack, 1953). Furthermore it has been shown that this turnover is largely confined to the small fractions of more soluble collagen that can be extracted from growing connective tissues (Harkness *et al.*, 1953).

The metabolism of chondroitin sulphate has been investigated with ^{35}S-labelled sulphate (Bostrom, 1952, 1953) and the results suggest that both in costal cartilage and in skin there is considerable metabolic activity in this fraction, the chondroitin having a half-life of about two weeks

Under conditions of limb atrophy (Slack, 1954) the uptake of labelled glycine by collagen is not reduced and is, if anything, greater than in normal limbs, which suggests that local atrophy of the fibrous framework is not due to a cessation of synthesis in a metabolic system of slow turnover but to some other mechanism, which presumably involves active removal.

Age Changes

Granulation tissue and the connective tissues of the young growing animal are relatively rich in polysaccharide, and contain a high proportion of argyrophil fibres; but with advancing years the proportion of polysaccharide becomes less and argyrophil fibres tend to be replaced by collagen fibres, so that the collagen polysaccharide ratio rises with advancing age (Sobel *et al* , 1953; Malmgren and Sylvén, 1952). As a result the connective tissues become less hydrated, less elastic and less resilient and therefore less able to resist mechanical forces; which is another way of saying that as old age creeps upon us we become shrivelled up and stringy and our limbs become stiff and readily damaged by trivial injuries.

Histologically defined fibrinoid has been investigated by X-ray diffraction and electron microscopy. As in all such studies the first requisite is an adequate sampling technique, but good correlation between histological appearances and biophysical findings has been achieved (Kellgren *et al.*, 1951). The results of these studies show that the parts of a rheumatoid nodule which show the histological appearances of normal collagen give an X-ray diagram of more or less disoriented collagen, and show plenty of normal collagen fibrils under the electron microscope. In the parts that on histological examination show advanced fibrinoid change the X-ray diagram typical of collagen may be entirely absent and is replaced by a diffuse ring with a spacing of about 4·8 Å. Material from such a spot, when examined under the electron microscope, may show only amorphous material and no collagen fibrils at all. Intermediate areas showing varying mixtures of normal collagen and abnormal material were found by all techniques, and in such intermediate zones the electron microscope revealed broken and disintegrated collagen fibrils as well as increasing amounts of amorphous material. It is interesting to note that all collagen fibrils seen during these electron microscope studies had the usual spacing of about 640 Å.

The precise chemical characterization of fibrinoid has not yet been achieved and indeed it may be very complex since the totally amorphous material is likely to differ from that which contains silver-staining fibres; but it is generally agreed that a fibrinoid area contains much material that resists boiling and has a low hydroxyproline (Ziff *et al.*, 1953) and a high tyrosine and hexosamine content and it seems likely that fibrinoid is related to the cement substance of normal connective tissue. The susceptibility of fibrinoid to various enzymes has also been studied (Glynn and Loewi, 1952) Some of this work and much of the chemical studies (Consden *et al.*, 1953) has been done on fibrinoid from patients with rheumatic fever, and it is not certain that this rheumatic fever fibrinoid is identical with the fibrinoid of rheumatoid disease. Nevertheless these studies illustrate the diverse ways in which an interesting connective tissue component can be investigated.

which accompanies experimentally induced hypertension (Wilson and Byrom, 1939) is possibly an example of this mechanism in subacute form. Conversely inadequate mechanical stimulation is associated with rapid deterioration and atrophy of the framework, so that an optimum play of mechanical force is required to keep the system in good condition, and the practice of orthopaedics is largely based upon this principle.

[illegible]

received very little attention but there is a whole group of rheumatic diseases which are characterized by disintegration of the supporting framework and the formation of fibrinoid in the connective tissues. Local disintegration of the framework also accompanies inflammation and is seen in its extreme form in pus-formation, but it is not certain that the disintegration which occurs in the rheumatic disease is always the result of inflammation. *Various mechanisms might lead to reduced stability.* Thus a stabilizer such as chondroitin sulphate might be degraded by enzymes such as hyaluronidase. Alternatively some antigenic cement might be destroyed by an antigen antibody reaction, or the protein component of the fibres might become degraded by enzymes like collagenase, but these are only hypothetical working models and the whole field of connective tissue stability in disease remains to be explored.

Thirdly there might be failure of synthesis. This should lead

[illegible]

do not heal. Thus gross malnutrition, scurvy and gross cortisone overdosage may be cited but, apart from this, failure of wound-healing is usually due to local causes such as defective blood supply or infection The question of an overactive removal mechanism is another possibility which is completely unexplored.

Then there may be failures of organization in which normal supporting material is laid down but in excessive or inadequate amounts; an example of the former would be keloid and of the latter osteogenesis imperfecta in which the whole supporting system, and not merely the bone, is inadequate. Faulty re-

The character of elastin also changes with advancing age (Lansing, 1951). It acquires a substantial content of calcium and phosphorus and an altered amino-acid composition. There is also a tendency for calcium to be deposited in tissues such as cartilage and tendon. Thus the various forms of dystrophic calcification may be exaggerated examples of the ageing process of either elastin or the collagen polysaccharide complex.

Lastly, the connective tissues are intimately involved in the processes of inflammation, but that is a whole subject on its own which was most admirably reviewed by a previous and more distinguished lecturer in this series (Cameron, 1952).

The System in Summary

In summary, the supporting system may be defined as the extracellular framework of the organism which in vertebrates is composed of a complex and varying arrangement of fibres which are largely protein but partly polysaccharide, set in a matrix which is largely polysaccharide but partly protein in nature. To perform its supporting and lubricating functions this system requires maximal physical stability and minimal metabolic activity. But it must be able to revert locally to a state of high metabolic activity in order to repair any damage caused by the operation of excessive mechanical forces. The cells of this system are presumably largely responsible for the initial production of the framework and for controlling its maintenance, remodelling and removal, but little is as yet known about the enzyme systems or other mechanisms by which these changes are achieved.

Disorders

Having defined the supporting system in this way, let us consider how it may fail to perform its functions.

Firstly the supporting framework may be broken down by the operation of excessive mechanical forces In a severe acute form this produces wounds, sprains and fractures in which as a rule the system is sufficiently reawakened to effect some degree of repair. In a less severe and more chronic form excessive mechanical stresses produce gradual disintegration resulting in disorders like degenerative joint disease. The arterial necrosis

CAMERON, G. R. (1952). *Lectures on the Scientific Basis of Medicine. Vol. I: 1951–52.* Pp. 130–42. The Athlone Press, London.
CHARNLEY, J. (1953). *Compression Arthrodesis.* E & S Livingstone, London.
CONSDEN, R., GLYNN, L. E. and STANIER, W. M. (1953). *Biochem. J.* **55**, 248.
CRUICKSHANK, B. and HILL, A. G. S. (1953). *Nature and Structure of Collagen* Edit. by J. T. Randall. Pp 27–32. Butterworth, London
DOLJANSKI, L. and ROULET, F. (1933). *Virch Arch.* **291**, 299.
EKHOLM, R. and NORBÄCK, B (1951). *Acta Orthop. Scand.* **21**, 81.
FELL, H. B. and MELLANBY, E. (1952). *J. Physiol.* **116**, 320.
GERSH, I. (1951). *Trans. 2nd Macy Conf on Connective Tissues.* Pp 10–43 New York.
GLYNN, L. E. and LOEWI, G. (1952). *J. Path. Bact.* **64**, 329.
HALL, D. A., REED, R. and TUNBRIDGE, R. E. (1952). *Nature, Lond* **170**, 264.
HARKNESS, R. D., MARKO, A. M., MUIR, H M and NEUBERGER, A. (1953) *Nature and Structure of Collagen.* Edit. by J T. Randall. Pp. 208–12. Butterworth, London.
[illegible]
[illegible] 52
KELLGREN, J. H., BALL, J , ASTBURY, W. T., REED, R and BEIGHTON, E. (1951). *Nature, Lond.* **168**, 493
KLEMPERER, P. (1950) *Am J. Path.* **26**, 505.
KLINGE, F. (1933). *Erg. Path* **27**, 1.
KRAMER, H. and LITTLE, K (1953). *Nature and Structure of Collagen.* Edit. by J. T. Randall Pp. 33–43 Butterworth, London.
LANSING, A I. (1951). *Trans. 2nd Macy Conf. on Connective Tissues* Pp 45–85. New York
[illegible]
[illegible] 88–100.
[illegible]
NEUBERGER, A , PERRONE, J. C. and SLACK, H. G. B (1951). *Biochem. J.* **49**, 199
NEUBERGER, A. and SLACK, H G B (1953) *Biochem J* **53**, 47.
NEUMANN, E. (1880). *Arch Mikr Anat.* **25**, 130.
PARTRIDGE, S. M (1948). *Biochem J* **54**, 387.
PORTER, K R and VANAMEE, P. (1949). *Proc Soc exp. Biol., N Y.* **71**, 513.
RANDALL, J T , FRASER, R D B., JACKSON, S., MARTIN, A V. W. and NORTH, A C T (1952). *Nature, Lond* **169**, 1029
RICHARDS, A G. (1951). *The Integument of Arthropods.* University of Minnesota Press, Minneapolis
ROLLETT, A (1871). *Handbuch der Lehre von den Geweben.* Edit. by Stricker. Pp 34–107

modelling as in experimental hypervitaminosis A (Mellanby, 1947) would fall in this category and many obscure dystrophies of bones and joints may be of this nature.

We should also consider alterations in connective tissue components, as for example dystrophic calcification or the Ehlers-Danlos syndrome in which collagen is replaced by elastin. Recent work has however shown that many of the so-called elastoses are not of this nature but result from alterations in collagen which enable it to accept stains which are only accepted by elastic fibres in normal tissues (Tunbridge *et al.*, 1952).

Lastly connective tissue components may be affected by the products of metabolic disorders in other systems. The ochronosis of alkaptonuria and the arthritis of gout are obvious examples; while the disorders of bone which result from disturbances of body electrolytes would also come in this category.

Conclusion

These examples suffice to illustrate the method of approach and need not be multiplied, especially since I have already indulged in more speculation than might seem proper in a lecture devoted to the scientific basis of medicine, but it has often been said that the most important step in scientific research is to ask the right question, and my purpose is to suggest that the development of new techniques makes it now profitable to pose a number of interesting questions about the biological properties and possible disorders of the supporting system.

VI

Hemispherectomy and the Localization of Function

E. ARNOLD CARMICHAEL

During the last hundred or more years a store of information has been built up regarding localization of function. To this physiologists, psychologists, neurologists and neurosurgeons have all contributed, and one might have expected that over so many years unanimity of opinion would have been reached as to the function of various parts of the human brain. It is true that there may be little conflict of opinion over the function of certain areas of the human brain—but one has only to turn to recent neurological literature to appreciate that divergence of opinion still remains. It is my purpose to present some observations which have been made during a study of patients who, suffering from a hemiplegia of early onset, have had the damaged cerebral hemisphere removed as a therapeutic measure designed to alleviate fits and behaviour disorders, concomitants of the hemiplegia. I will refer to observations made in association with Dr. Bates and to work carried out by Mr. McKissock and others. I readily and gratefully acknowledge their kindness in permitting me to avail myself of their material.

The observations may be grouped under several headings: first, the disturbances of movement resulting from the primary lesion to the brain and how such movements are affected by removal of the damaged hemisphere. secondly, the results of electrical stimulation of the brain carried out by Dr. Bates during the surgical removal of the hemisphere; thirdly the

SLACK, H. G. B. (1954). *Clin. Sci.* **13**, 155.
SOBEL, H., ZUTRAUM, H. A. and MARMORSTON, J. (1953). *Arch. Biochem Biophys.* **46**, 221.
TUNBRIDGE, R. E., TATTERSALL, R. N., HALL, D. A., ASTBURY, W. T. and REED, R. (1952). *Clin. Sci.* **11**, 315.
WILSON, C. and BYROM, F. B. (1939). *Lancet*, **i**, 136.
WOLBACH, S. B. (1933). *Am. J. Path.* **9**, 689.
WOODIN, A. M. (1952). *Biochem. J.* **51**, 319.
ZIFF, M., KANTOR, T., BIEN, E. and SMITH, A. (1953). *J. Clin. Invest.* **32**, 1253

repeated focal fits ushers in a complete hemiplegia, from which only partial recovery may take place. In this last group, the evidence available suggests that the pathological process is commonly maximal in one hemisphere, but there are grounds for believing that in many the other hemisphere may be affected, only to a much less degree.

Motor Functions

I will now consider the defects of movement which may be observed in these subjects. In those children with a prenatal vascular abnormality of greater severity in the occipital lobe than in the parietal or frontal lobes, the parents have stated that, though the facial naevus has been obvious at birth, no defect has been noted in the movements of the limbs of the two sides up to the age of three months. At this age the child commences to move the limbs of one side, that opposite to the facial naevus, rather less than the other. It may be that no defect is noted till a later age—nine months—at about the time the child attempts to support itself or move about in the cot. It is clear that though the pathological lesion is a congenital one the defect, if any, in movement is not readily observed till the child is three months old; the defect then observed is a deficiency in movement—not a total loss. In patients with a lesion limited to the territory supplied by the middle cerebral artery—and here I would offer reminder that the area includes the lower half or two-thirds of the precentral and postcentral gyri, as well as the insula and upper temporal gyrus—the time at which a motor deficit has been observed has varied from patient to patient. The common story is that no inequality in amount, or difference in character of movement of the limbs on the two sides has been noted till the third or fourth month of life; sometimes, however, the parents have observed nothing abnormal till the child reached six or nine months We ourselves observed one infant from shortly after birth, and it was not till the age of [illegible]

[illegible]

of the limbs of one side has followed upon an episode of status epilepticus at an age of one or more years, up till when develop-

effect of hemispherectomy on sensory functions; and lastly, the pattern of the epileptic attacks before and after operation. At this operation Mr. McKissock removes the frontal, parietal and occipital lobes: only a small portion of the hippocampal gyrus is left after the removal of the temporal lobe. A part of the globus pallidus with putamen, external capsule and insula is removed. The thalamus, however, is left intact. All that remains to fill the cavity is the stump of the basal ganglia with a small portion of the hippocampal gyrus. In the twenty-five patients who have undergone this operation, three types of lesion have been found: first a condition, probably congenital, in which one hemisphere is involved by a vascular abnormality associated with calcification affecting the cortical grey matter, most extensively in the occipital region, but spreading forward into the parietal and temporal lobes and sometimes into the frontal lobes; an ipsilateral facial naevus is also present at birth. There is the second group, in which the territory of the brain supplied by the middle cerebral artery is replaced by a cyst commonly communicating with the ventricle; this lesion, probably an occlusive one of the artery, occurs as far as is known at or about the time of birth. I have been privileged, through the kindness of Professor Linnell of Toronto, to see the brain of an infant who died in the first two days of life; a very recent embolus blocked one middle cerebral artery. Our patients with a 'cystic' lesion have suffered from disturbance of respiration at birth, though commonly nothing abnormal has been noted until the third or fourth month of life. In one child, aged four months and not using the right arm and leg as much as the left limbs, an air-encephalogram taken within a few days of this observation showed cystic formation of the brain in the distribution of the left middle cerebral artery. Such cysts are known to take several weeks to develop; accordingly the infarction in this infant must have occurred some time before the poverty of movement had been noted. So far as can be demonstrated by clinical investigation, air-encephalography and arteriography, the lesion in these two groups is limited to one hemisphere. There is a third group, called 'diffuse', in which after a period of normal development over one, two or more years, a sudden episode characterized by

impressed on the affected limbs, with coincident restriction of movement, probably arises as a result of the maturation of structures in the brain-stem. I have been fortunate, through the kindness of Dr. Gooddy, to see a boy aged three years, who until three months old appeared to develop as a perfectly healthy child. The mother then noticed that the child regressed—movements became less and it was established that the child was failing to react to visual stimuli. Gradually the child developed a posture in which both upper limbs were flexed with the legs extended. When the child died at the age of three years, it was found to have a cerebral mantle with the configuration of a foetus aged six months—the cortico-spinal tracts were not myelinated yet the hypothalamus, brain-stem and cerebellum all appeared normal. At the sixth month of pregnancy the mother of this child had a ruptured ovarian cyst removed. In this child, in whom the cerebral cortex had not developed beyond that of a six months foetus, but in whom the brain-stem had, an abnormal and limited pattern of limb movement was noted only when the child reached its third month of life. If we accept the inference from the second case that the developing brain-stem in the absence of maturing cerebral hemispheres imposes a posture with consequent degradation of movement, it would appear reasonable to suggest that the increasing poverty of movement seen to develop in a child with a damaged cerebral hemisphere is the expression of the function of a maturing brain-stem.

Infants with damage to the region supplied by the middle cerebral artery are never able as they develop to carry out those movements which permit the use of the affected limbs in any skilled activity There is poverty of movement of the affected limbs, a deficiency which is greater in the digits. It is seldom that an entire absence of movement of the digits is observed; commonly the digits move in unison, though movement of an individual digit alone may occasionally be seen. There is, in other words, no absolute paralysis.

In all three types of lesion there are areas within the damaged hemisphere which on histological examination appear to have a cellular configuration within normal limits; Nissl staining

ment has been normal. After the episode, a complete motor paralysis of the limbs of one side has been noted, which over weeks or months has shown some recovery.

Why is it that a prenatal lesion, and one occurring at about the time of birth, shows little or no manifestation of its presence until about the third month of life? What part, if any, is played by the cerebral hemisphere in the production of movement observed during the earliest weeks of life? Through the kindness of Dr. Martin and Dr. Bodian I have seen a film of a child aged a few weeks in which the infant moved both legs and arms and took nourishment from a bottle. A few days after this film had been taken the child died, and Dr. Bodian found post-mortem an absence of the cerebral mantle; above the tentorium only the stumps of the basal ganglia and a small part of the temporal lobes were present. This is no isolated case, for scattered throughout neurological literature similar cases are reported. It is difficult to escape the conclusion that the movements to be observed in an infant at birth and for some weeks after are possible without a cerebral cortex. An occlusive lesion at about the time of birth involving the middle cerebral artery and causing damage to the so-called motor cortex need not be associated with poverty of movement during the first few weeks of life. The pattern of movement seen in the early weeks of life is thus apparently largely determined by the function of levels lower than those of the cerebral cortex. Indeed it may be essentially a spinal and primitive brain-stem pattern.

Concurrent with the increase in range and dexterity of movement of an infant's limbs and especially of the hand, there is a progressive development of the cerebral hemispheres. I need only call to mind the steady process of myelination and the changes in the distribution and character of the cortical nerve cells. These changes of maturation take place not only in the cerebral hemisphere but also in the brain-stem. It is of interest that in infants with a natal hemispheric lesion, as the healthy limbs show an increase in range and skill of movement, the affected limbs show a progressive depletion of movement associated with the adoption of a posture, flexion of the arm and fingers with extension of the leg. This posture, which is being

bute significantly to the pattern of movement observed in either the upper or lower limb.

An important function in which motor elements play a considerable part is speech. Our series is too small to permit of any useful analysis. It may, however, be of interest to note that removal of the right or left hemisphere has not added to any defect of speech which may have been previously present, except in one patient. In this patient, with a lesion of the left hemisphere occurring during the first two years of life and in whom speech was possible before the operation, speech has ceased since the left hemisphere has been removed. There are on the other hand a number of patients who, following the removal of the hemisphere, have exhibited an improvement in the capacity to speak.

SENSORY FUNCTIONS

Let me now pass to some observations in the field of sensory function. The ready distractability exhibited by so many of these patients has meant that elaborate, careful and repeated testing has had to be limited to only a few. There have been four patients, each of sufficient intellectual capacity and age to appreciate the care and attention needed, and on whom satisfactory observations have been possible. Three suffered from an occlusive lesion of the middle cerebral artery at or about the time of birth, and the remaining one had an acute episode at one year, following which she developed a hemiplegia. Three have had a left hemiplegia and one a right hemiplegia. The results of our observations have been essentially the same in all four patients The threshold for light touch has been slightly raised on the hemiplegic side, especially on the lips and fingers, while the quality of sensation experienced during stroking the affected side with a wisp of wool has lacked the usual tickling element The threshold to pinprick has been the same on both

sides, but there has been frequently a hesitancy before replying

shows normal lamination with apparently healthy cells, and myelin stains show apparently healthy nerve fibres. In occlusive lesions of the middle cerebral artery the structures in the upper third of the precentral and postcentral gyri are apparently histologically healthy. Has this apparently undamaged tissue any obvious function? Bates and McKissock (1951) have shown that electrical stimulation of this apparently healthy cortex may result in movement of the contralateral limbs. The movement following stimulation closely resembles in pattern that movement observed during the clinical examination of the patient. Recently Bates (1953b) has reported that electrical stimulation of the exposed internal capsule after the pathological hemisphere has been removed results in a similar pattern of movement in the affected limbs to that obtained from stimulation of the cortex. His evidence is that this apparently healthy cortical tissue has a functional connection with spinal levels.

Does removal of this remaining and apparently healthy tissue, which on stimulation results in movement of the contralateral and affected limbs, add in any way to the motor deficit already observed in these patients, or change the pattern of movement? Dr. Bates and I have sought diligently for evidence indicating an increased impoverishment of movement. In those whose ablated hemispheres have had small scattered lesions throughout them we have detected an increase in the deficit of movement. In the cystic group, however, we have seldom observed any addition to the disability; in fact, the pattern of movement of the fingers has remained unaltered following hemispherectomy, and in some the usefulness of the hand as an instrument for holding objects such as a fork or spoon has even improved In nearly all, movement of the hemiplegic leg has improved and our records contain frequent reference to 'being able to walk better', to 'now being able to dance' and to a reduction in the wearing down of the shoe of the hemiplegic limb. It is thus clear that removal of further and apparently healthy portions of the so-called motor cortex is not necessarily associated with an increased deficit of movement of the hand; in some it may be accompanied by an improvement in the usefulness of the leg. The remaining undamaged cortex does not apparently contri-

though movement of the hand was possible. The stimulus of putting something into the right hand did not cause a motor response. When last seen many weeks following the hemispherectomy, it was clear that this had changed and the affected hand was now being used to hold objects. In the infant who had no cerebral hemispheres, touching its arms caused the child to [illegible]

cerebral hemispheres had the configuration of a six months foetus, sensory stimulation provoked movement—pinprick was succeeded by movement of the limbs stimulated. The complicated movements of chewing followed the inhalation of odours, and if moved or disturbed displeasure was indicated by crying; sensory stimulation thus provoked quite complex movements. Movement of a complex nature may therefore be provoked when spino-cortical or thalamo-cortical projections are defective In patients with an infantile hemiplegia, both before and after hemidecortication, the sensory defect is essentially a failure of the discriminating sensory functions. 'Skilled' sensation is as defective as is 'skilled' movement.

The Remaining Hemisphere

The disturbances of function discussed so far were determined by comparing the efficiency of the affected half of the body with the apparently healthy half. By using this technique of comparison, we have assumed that the functions of the healthy half are unaffected by damage to or removal of the cerebral hemisphere on the same side. Is there any reason to believe that there may be some functional defect of the ipsilateral limbs? It is known from the work of Penfield and Welch (1951) that cortical stimulation may provoke bilateral sensation and movement. Bates (1953a) has shown that electrical stimulation of the mesial surface of the remaining hemisphere after hemispherectomy produces movements of the ipsilateral limbs as well as of the contralateral limbs. Aware of this, both Dr. Bates and myself have been on the constant look-out for any evidence of a deficit of motor or sensory function on the good half of the body.

when the stimulus was applied to the affected half of the body. *Sense of movement at a joint has been only slightly impaired* though the patient may be quite unable to detect the direction of movement at a finger or toe joint of the affected limbs. Discrimination between one and two points has been defective over the affected half of the body, most marked on the lips and fingers. Recognition of objects in the affected hand has always been very poor. Location of the affected limb in space is never accurate. Again, location of a stimulus applied to the hand or forearm is at fault, being frequently displaced proximally. *Hemispherectomy has caused little or no change in these sensory* disturbances. In two patients the threshold for a painful stimulus has been definitely raised. Careful questioning has never elicited the expression that either before or after the removal of the hemisphere has the patient been unaware of the affected limbs.

While such detailed observations were possible in the more co-operative and older patients, it was noted in the younger children, especially those under the age of ten, that the awareness of stimuli applied to the affected limb seemed less acute. Frequently it has been possible to touch the hemiplegic arm without the child giving any indication that anything had occurred —yet a similar stimulus to the healthy limb provoked an aggressive gesture. After removal of the hemisphere when the child moved about the ward it behaved as if unaware of the affected limbs. The child would walk into the room, thrusting the good limbs forward with the body rotated so that the affected limbs were always behind the good limbs. the affected arm or hand might become entangled in the clothes during dressing, but no attempt would be made to disengage it. If the child was invited to throw a box placed in both its hands, it would often not notice that the release of the box was hindered by the affected hand still retaining it. Over weeks, many of these signs would disappear, so that eventually the child's state was back to its pre-hemispherectomized level. In the youngest infant whose hemisphere was removed at the age of nineteen months, the right and affected arm was never used immediately following the operation to take hold of toys placed in that hand,

have been noted to occur. It has been established from cinematograph records that the first movement may occur in the affected limb, when the cortico-spinal pathway from the opposite hemisphere to that limb has been divided in the region of the internal capsule, and the cortical cells of origin of that pathway entirely removed. Though Bates (1953a) has reported that following electrical stimulation of the mesial surface of one hemisphere, after the other hemisphere had been completely removed, movements occurred in the ipsilateral and hemiplegic limbs, as well as in the contralateral and sound limbs, I hesitate to believe that these epileptic movements in the hemiplegic limbs are the direct result of a discharging lesion on the mesial surface of the good hemisphere.

Summary

I have drawn attention to the site of the lesion for which the operation of hemispherectomy has been carried out and shown you that it has involved the so-called motor and sensory cortex. The lesion may occur during the period of development of the cerebral hemisphere, and affects function which may be expressed in terms of poverty of movement. There has also been a loss of sense of direction of movement and of discriminatory ability, and I have dared to call the sensory deficit that of 'skilled' sensation. These functions of discrimination and skilled movement are apparently dependent upon the full development of the cerebral hemisphere after birth. A movement of skill is essentially a sensory-motor function dependent on the full integrity of the cerebral hemispheres. The pattern of a skilled movement is imposed upon the pattern of movement originally set by spinal and primitive brain-stem structures. Evidence has been provided which may be taken as indicating that one cerebral hemisphere may be concerned, not only with movement of the limbs of the opposite side of the body, but also, though to a lesser extent, with movements of the limbs on the same side of the body.

Neither following early damage to the one hemisphere nor after the removal of that hemisphere have we been able to satisfy ourselves that disturbance of function occurs on the ipsilateral half of the body. It may be that our tests have not been appropriately designed to determine dysfunction, but dysfunction we have not detected.

Fits

One of the purposes of the operation has been to stop the fits and in this it has been remarkably successful in the majority of the patients. There have been, however, a few patients who have continued to have fits. The character of the fits subsequent to the removal of the hemisphere provides interest. In two subjects, the fits of which they originally made complaint were ushered in by a sensory experience. In one the first indication of a fit was a sensation of 'something running up the veins' of the affected forearm, which was followed by a loss of consciousness and movements of all four limbs. Since the damaged hemisphere has been removed the patient has had eight attacks in each of which she has experienced for a few seconds only an exactly similar sensation in the affected arm followed by a slight disturbance of full awareness but not by unconsciousness nor by movement. In the other patient an indescribable sensation located in the weak forearm and passing up to the shoulder—'almost a pain'—preceded the loss of consciousness and the motor element of the fit. Since the operation he has on several occasions experienced the sensation in the weak limb and has called the nursing staff, believing he might be going to have a fit—yet no loss of consciousness nor movement of any limb has occurred. In both these patients, an apparently strictly localized momentary and subjective sensory phenomenon has occurred in the weak arm after the opposite cerebral hemisphere has been removed. One can only conjecture as to the site of the discharging lesion responsible for such sensations. The thalamus remains Could the discharge occur within its territory, or has it taken place in the remaining hemisphere or even at a lower level? We have no means of determining this In other subjects, fits have been observed after the operation; and in these, definite movements in all four limbs

VII

Anticholinesterases

W. D. M. PATON

ANY account of the anticholinesterase drugs must begin by taking into consideration the properties and the distribution of the cholinesterase enzymes as well as of the drugs which inhibit them. This immediately leads us into ground so complicated that one is in danger of losing the original and rather straightforward background to these agents. It is therefore worth while to outline this basic story and to consider the later developments as elaborations or divergencies from it.

The preliminary rumblings are heard in Fraser's study of the toxicity of the esere bean in 1860–70, and of the antagonism of atropine to it; we now know that the Calabar natives in their choice of an ordeal poison had anticipated modern chemical warfare by many years. Then comes the observation by Dale in 1914 *of the brevity of the action of acetylcholine, prompting his* suggestion that an enzyme exists in the body able to destroy acetylcholine. Loewi and Navratil (1926) showed that 'Vagusstoff' (later to be identified as acetylcholine) could be rapidly [illegible]

his colleagues, amongst them neostigmine. It is interesting that they were developed primarily as miotics and were only later recognized as anticholinesterases. From 1934 onwards came the *fundamental experiments of chemical transmission*. These elucidated how eserine produces its typical effects, by preserving acetylcholine released at nerve-endings, so favouring trans-

REFERENCES

BATES, J. A. V. (1953a). *Brain*, **76**, 405.
BATES, J. A. V. (1953b). *J. Physiol.* **123**, 49P.
BATES, J. A. V. and MCKISSOCK, W. (1951). *J. Physiol.* **115**, 51P.
PENFIELD, W. and WELCH, K. (1951). *Arch. Neurol. Psychiat., Chicago*, **66**, 289.

that they named it 'true' cholinesterase, while the second was only apparently concerned with this and hence was named 'pseudo'-cholinesterase. They further showed that one could measure the true cholinesterase in a tissue or extract by the rate of hydrolysis of acetyl-β-methyl-choline (a substance chemically very like acetylcholine), and the pseudo-cholinesterase by the rate of hydrolysis of benzoylcholine (a rather different compound). This christening was soon challenged, partly because words like 'true' and 'pseudo' had an air of moral judgement, and partly because it was found that some cholinesterases existed whose properties diverged from those laid down. But these names reflect a useful physiological approach, and though they may be criticized, their use is convenient, nearly self-explanatory and not fundamentally misleading, since the conception implied has been generally accepted, even if the names have not.

Other workers found that the 'true' cholinesterases could hydrolyse a number of substances possessing acetyl groups but not choline groups, so that they thought that a better name for the true enzyme would be 'acetylase'. A third group of workers analysed the kinetics and specificity of the enzymes, stressing (1) the high rate of reaction of the true enzyme with acetylcholine, (2) the fact that it seemed designed to work most efficiently in low dilution, and (3) the fact that it seemed to be fairly particular about the resemblance of its substrates to acetylcholine They painted a picture whereby true cholinesterase seemed to bear a highly specific relationship to acetylcholine while the pseudo-enzyme was not *specifically* concerned with it; hence the names 'specific' and 'non-specific' cholinesterase. A fourth approach arose from experiments varying the aliphatic chain in acetylcholine This revealed that 'true' split acetyl esters more rapidly than any other ester, while the 'pseudo' split butyric esters much better. Hence the names 'aceto-cholinesterases' and 'butyro-cholinesterases'. A fifth approach is to abandon attempts to give an explanatory name and simply refer to them as 'Group I' and 'Group II' enzymes, or 'α' and 'β' enzymes, names which could obviously be retained whatever properties were revealed about the enzymes. Another approach

mission, unless the accumulation of transmitter went to excess and began itself to cause synaptic block.

Given these facts of cholinergic transmission and appreciating the importance of an enzyme for destroying acetylcholine once released (to allow independent repeated shocks, and to prevent block by accumulation) then one might expect to describe the situation in the following simple terms:

(1) there would be a single enzyme whose action is specific to acetylcholine and which acts on acetylcholine with great rapidity and in high dilution;

(2) this enzyme would be localized at the sites of acetylcholine action only:

(3) the enzyme would be specifically antagonized by a number of inhibitors which resemble the transmitter chemically;

(4) the effect of these inhibitors of the enzyme would all be due to the accumulation of unhydrolysed acetylcholine and the intensity of their effects would be in proportion to the degree of enzyme inhibition.

As a framework for the discussion we shall take each of these simple statements in turn, and see how far they have come to need expansion or qualification.

The Enzymes Destroying Acetylcholine

It was quite early recognized that more than one enzyme in the body could destroy acetylcholine. Since these observations were made, a considerable literature has grown up, a good deal of which is concerned with what the enzymes should be called (for reviews see Augustinsson (1942) and Whittaker (1951)). This has tended to make the subject rather confusing, but it can be simplified somewhat if one relates each proposed nomenclature to the particular point of interest of the investigator concerned. First was Mendel and Rudney, who studied two crude enzymes both of which could at first split aliphatic esters as well as choline esters. When they purified these preparations, however, they found that one ceased to hydrolyse the aliphatic esters, while the other continued to do so. On the basis of this and other evidence they concluded that the first enzyme was really and truly concerned with hydrolysing acetylcholine so

also be identified in the cholinergic electric organ of *Electrophorus*, in cholinergic nerves, in autonomic ganglia (chiefly but not only localized in the preganglionic nerve terminals), in the adrenal medulla, in the central nervous system (particularly the basal nuclei) as well as in ganglia in the intestine. These sites are reasonable; within nervous tissue, particularly in the brain, its presence can be taken as a sign (along with choline-acetylase activity, and the presence of acetylcholine itself) of cholinergic activity by the nerves in which it is found. But it is less easy to understand why an enzyme conforming accurately to true cholinesterase should also exist in two nerve-free tissues, the placenta and the red cell, as well as in such unexpected materials as cobra venom (Koelle, 1951; Whittaker, 1951).

Pseudo-cholinesterase has a wider, indeed almost universal, distribution.

Our simple picture of an enzyme being localized only at the sites of acetylcholine released must therefore be modified; and we have instead to think of 'true' cholinesterase being localized at cholinergic synapses in the way expected, but also of true cholinesterase also occurring in other places, and of pseudo-cholinesterase being distributed very widely indeed (for reasons still not understood). A number of suggestions about the role of the latter can be made, if, as Professor Burn has suggested, acetylcholine is concerned in spontaneous rhythmic activity of tissues such as heart muscle, then pseudo-cholinesterase might have some important non-synaptic role there. Or secondly, it may be that pseudo-cholinesterase 'backs up' the true enzyme, particularly if (as Koelle (1953) points out) it can help the latter when high concentrations of acetylcholine occur. But it is much too early to discuss this in detail.

The Inhibitors of Cholinesterase

A useful way of summarizing these is to relate pharmacological action to their structural resemblance to acetylcholine (Fig. 1). It is now believed (largely from Nachmanssohn, Wilson and Bergmann's work on *Electrophorus* esterase) that the molecule of acetylcholine reacts with the enzyme at least at two sites: an anionic site, negatively charged, which associates with the

of the same sort was simply to state where the enzyme came from or at least where the prototype enzyme came from. Thus the 'true' enzyme was called 'E-cholinesterase' for erythrocyte cholinesterase and the 'pseudo' was called 'S-cholinesterase' for serum cholinesterase.

One is reminded of the White Knight in *Through the Looking Glass* (Carroll, 1878), where he explains to Alice that the song he is about to sing has a name *called* 'Haddocks Eyes', although its name *really is* 'The Aged, Aged Man', but the *song* is called 'Ways and Means' and the song *really is* 'Sitting on a Gate'. The names given to the enzymes splitting cholinesterase could well be regarded in the same light.

Significant points from all this work could be summarized as follows. First, there is one enzyme (or possibly a group of very closely related enzymes) which (*a*) acts only on esters closely related to acetylcholine such as acetyl-β-methyl-choline; (*b*) acts very rapidly on acetylcholine; and (*c*) is favoured over an important range of concentration by dilution of acetylcholine (so that it actually comes to work less effectively when the concentration of acetylcholine rises). Second, there is also another and rather more puzzling family of pseudo-cholinesterases which split acetylcholine less rapidly, do it more efficiently the more concentrated the acetylcholine (thus working relatively inefficiently at low dilutions), and typically hydrolyse benzylcholine and butyric esters, although different members of the family differ somewhat in their favourite food. These facts do not matter much clinically, but they are fundamental in many investigations when one is trying to find out on what enzyme an inhibitor is acting, when it produces its characteristic effects.

The Distribution of Cholinesterases

Until recently acetylcholine's function in the body was viewed as being concerned solely with the transmission of activity across the various synapses, where, in fact, cholinesterase can be found. The most beautiful example is that of cholinesterase at the motor endplate, where it has been shown by a number of histochemical procedures (Koelle, 1951; Denz, 1953). It can

by possessing a quaternary head. Here we must recall that this head seems to confer an ability to stimulate the endplate (seen outstandingly in decamethonium and succinylcholine) or sometimes to paralyse it. Therefore we are not surprised to find that it is a powerful antiesterase, but that it can also exert a decamethonium-like and, perhaps, with long-continued dosage, a curare-like action at the endplate. Notice also that it is quaternary; this means that it cannot cross the blood-brain barrier easily and so is relatively free of the central effects exerted by eserine on the one hand and fluorophosphates on the other.

Third is *Tensilon*; this looks definitely inferior so far as equipment for esteratic competition goes: and in fact it is not a highly active anticholinesterase. It has, as you would expect, some muscle stimulant action, and some curare action under big doses and the right conditions. But a new element appears here —that of sensitization of the endplate independent of antiesterase action: for instance, it favours decamethonium action quite strongly (Paton, 1951). As a consequence, its anticurare action is out of all proportion to its anticholinesterase action. Both Pelikan, Smith and Unna (1954) and Hobbiger (1952) maintain, however, that its antiesterase action accounts fully for its anticurare action: perhaps one had better suspend judgement on this particular point at present. Be that as it may, Tensilon differs in having even more of a direct muscular action than neostigmine, and in having a notably quicker and more transient action (perhaps because it is a smaller molecule), but it resembles neostigmine in being free of central action.

An amazing product is the latest compound, 3113CT, which consists of two neostigmine molecules linked by an aliphatic chain (Funke, Depierre and Krucker, 1952). It can inhibit dog's red-cell cholinesterase at concentrations of 10^{-14} to 10^{-15} (compared to eserine 10^{-8}), and is able to produce muscarinic signs of cholinesterase inhibition with a dose of a few micrograms per kilo. It will be fascinating to see what develops from this

Now all the compounds so far described are nitrogenous bases, competing reversibly with acetylcholine. The last category, developed partly for military purposes, partly as insecti-

CHOLINESTERASE

NAME OF INHIBITOR — ANIONIC SITE — ESTERATIC SITE

Acetylcholine

Eserine (Physostigmine)

Neostigmine

Tension

Phosphorus compounds

Decamethonium

Fig 1

quaternary head, and an esteratic site, at which the split takes place, with an affinity for the ester arrangement of atoms.

If we consider *eserine* first, we see that it fits fairly well into both sites, while not having identity with acetylcholine at either position. It is therefore well-equipped to *compete* with acetylcholine for the receptor sites, and (incidentally) also to be itself slowly hydrolysed by the enzyme.

Neostigmine is in one structural respect closer to acetylcholine,

doses of the other. Returning to our earlier metaphor, one can say that an esteratic flirtation will protect against a fatal passion!

The phosphorus inhibitors vary, of course, among themselves in several respects. *First*, their lipoid solubility; this is greatest for DFP, then paraoxon, then TEPP and eserine, then OMPA and neostigmine. This corresponds roughly to their liability to cause central effects (Burgen and Chipman, 1952). *Secondly*, their duration of action: some data on the duration of meiosis after application to the eye serve to illustrate this: the effects of DFP lasted 450 hours, TEPP 190, eserine 80, neostigmine 30. *Thirdly*, some of them proved to be much less active *in vitro* than *in vivo*. This is because such compounds are metabolized in the liver to a more active (although less stable) product; this has now been shown for Parathion (Aldridge and Davison, 1952), and OMPA (Dubois *et al.*, 1951; Gardiner and Kilby, 1952). It has even been shown that removing the liver can protect an animal against several lethal doses of OMPA (Cheng, 1951). The others, such as DFP, TEPP, and the bases, have roughly proportionate actions under all conditions. *Fourthly*, different inhibitors vary somewhat in their proportionate antagonism to true and pseudo-esterases. DFP for instance can inhibit particular pseudo-esterases in about 200 times smaller concentration than that needed to inhibit a true enzyme (v. Austin and Berry, 1953). This is often helpful in studying the enzymes, and it explains, or helps to explain, why DFP sometimes is less active than expected; one can suppose that it is first taken up and fixed by the plasma enzyme, without having a chance to get at true enzyme localized at more strategic points. There is, unfortunately, no great regularity of behaviour in these proportionate antagonisms, and sometimes the discrimination between pseudo and true is much slighter. Conversely, some compounds interfere with true cholinesterase rather more readily, particularly the quaternary salts 284C51, and Nu 1250 (Austin and Berry, 1953); in fact decamethonium has some action of this sort (Paton and Zaimis, 1949) It may be possible, ultimately, to exploit these differences: but of course quaternary salts tend to have a good many

cides, a very wide one which is condensed enormously in the table, is of quite a different kind usually termed 'irreversible'. They are all compounds of organic phosphorus, and are believed to act by splitting off phosphonium ion ($O{=}P\langle^{R_1}_{R_2}$) which then phosphorylates the esteratic site. One can represent the situation by comparing cholinesterase to an attractive girl at a dance who usually breaks (hydrolyses) all hearts. Then one evening she finds a partner with a great mutual affinity (say eserine) who temporarily stops her multiple flirtations, by dancing with her all evening (his own heart being only broken a little bit); but when the dance is over, they separate and go home—the inhibition was reversible. With the phosphorus compounds it is quite different, and more dramatic, a suicide pact, in which both the partners are destroyed for ever. With DFP for instance, it seems that the enzyme has to be regenerated before activity returns, although with some of its relatives, such as TEPP and Paraoxon, the affair is reversible in its early stages.

This tendency to irreversibility is the first characteristic of the phosphorus compounds—bringing with it, of course, a much longer duration of action. But we can also predict (as is the case) that these drugs will be free from direct muscular actions—because they lack a quaternary group entirely. Further, since they lack the highly soluble nitrogenous group, they are liable to be fat soluble, sometimes actually volatile, and hence well absorbed into the lungs or from the skin, the gut, the eyes, and able to penetrate into and be taken up in the brain. They tend also to be 'fixed' in the tissues: for instance neostigmine injected into an arm exerts systemic effects, but DFP, in a comparable or larger dose, does not (Harvey *et al.*, 1947). There is one last important point: one might ask what happens when you combine the two types of anticholinesterase. If you give a phosphorus compound, then eserine or neostigmine subsequently simply adds on to it a little. But if you give eserine or neostigmine first, it seems to prevent the uptake of the irreversible inhibitor; and one gets the interesting situation that a small dose of reversible inhibitor will actually protect against several lethal

doses of the other. Returning to our earlier metaphor, one can say that an esteratic flirtation will protect against a fatal passion!

The phosphorus inhibitors vary, of course, among themselves in several respects. *First*, their lipoid solubility; this is greatest for DFP, then paraoxon, then TEPP and eserine, then OMPA and neostigmine. This corresponds roughly to their liability to cause central effects (Burgen and Chipman, 1952). *Secondly*, their duration of action: some data on the duration of meiosis after application to the eye serve to illustrate this: the effects of DFP lasted 450 hours, TEPP 190, eserine 80, neostigmine 30. *Thirdly*, some of them proved to be much less active *in vitro* than *in vivo*. This is because such compounds are metabolized in the liver to a more active (although less stable) product; this has now been shown for Parathion (Aldridge and Davison, 1952), and OMPA (Dubois *et al.*, 1951; Gardiner and Kilby, 1952). It has even been shown that removing the liver can protect an animal against several lethal doses of OMPA (Cheng, 1951). The others, such as DFP, TEPP, and the bases, have roughly proportionate actions under all conditions. *Fourthly*, different inhibitors vary somewhat in their proportionate antagonism to true and pseudo-esterases. DFP for instance can inhibit particular pseudo-esterases in about 200 times smaller concentration than that needed to inhibit a true enzyme (v. Austin and Berry, 1953). This is often helpful in studying the enzymes, and it explains, or helps to explain, why DFP sometimes is less active than expected; one can suppose that it is first taken up and fixed by the plasma enzyme, without having a chance to get at true enzyme localized at more strategic points. There is, unfortunately, no great regularity of behaviour in these proportionate antagonisms, and sometimes the discrimination between pseudo and true is much slighter. Conversely, some compounds interfere with true cholinesterase rather more readily, particularly the quaternary salts 284C51, and Nu 1250 (Austin and Berry, 1953); in fact decamethonium has some action of this sort (Paton and Zaimis, 1949). It may be possible, ultimately, to exploit these differences: but of course quaternary salts tend to have a good many

other actions in addition to antiesterase effects, which have complicated this sort of progress. *Fifth*, it must be mentioned that the phosphorus compounds do not only inhibit cholinesterases but also some other enzymes, the so-called aliesterase, the proteolytic enzyme chymotrypsin, and the B-type esterase of mammalian sera (Mendel, Myers, *et al.*, 1953).

Our simple picture, therefore, of reversible specific inhibitors chemically related to acetylcholine can well be retained, provided we remember the two types of attachment to the enzyme, and allow for a combination of inhibitor with receptor which is at least so slowly reversible as to be, for practical purposes, permanent. There are, however, important complications: most inhibitors are pretty active against both enzymes, a few much more against one than the other; and cholinesterase is not the only type of enzyme inhibited.

There is one inhibitor of cholinesterase particularly active against pseudo-esterase, which comes in a category by itself, viz. triorthocresylphosphate (TOCP). This has been known for some years to produce a central nervous paralysis; it was later found to be an inhibitor of cholinesterase and later still found to be particularly active against pseudo-cholinesterase in the brain (Earl and Thompson, 1952). The interest in it arises from the fact that it produces central nervous system lesions of a kind also seen with two of the anticholinesterases, which will be discussed later.

Effects of Inhibition of Cholinesterase

It is fairly straightforward to interpret most of the effects seen in the terms of our original simple picture as being due to accumulation of acetylcholine, so that the results of inhibiting cholinesterases in the body provide a concise exercise in the anatomy of cholinergic nerves. The careful studies by Harvey and his colleagues on the effect of DFP on man, still provide one of the best descriptions of these effects. (Table 1.) The most prominent effects are the muscarinic ones. Of the nicotinic actions of acetylcholine only that on striated muscle reveals itself at all obviously. Remaining is a group of central nervous system actions, some of a quite remarkable kind.

There has been a good deal of discussion as to whether the

TABLE 1. Symptoms Following the Daily Intramuscular Administration of DFP
(From D. Crob, J. L. Lilienthal, A. M. Harvey and B. F. Jones (1947) *Bull. Johns Hopk. Hosp.* 81, 217–44)

Effector organ	Symptom	No. of subjects (Total: 50 normal and 10 myasthenic)
1. Muscarinic		
(*a*) Gastro-intestinal		41
	1. anorexia-nausea	29
	2. abdominal cramps	25
	3. vomiting	14
	4. diarrhoea	12
	5. cardiospasm	4
	6. nausea after smoking	4
(*b*) Sweat glands	increased sweating	10
(*c*) Lacrymal glands	increased lacrymation	3
(*d*) Salivary glands	increased salivation	2
(*e*) Pupils	slight miosis	2
(*f*) Ciliary body	difficulty of distant vision	2
(*g*) Lungs	respiratory difficulty, suggestive of broncho-constriction	2
(*h*) Bladder	urinary frequency	2
(*i*) Heart	slight bradycardia with premature contractions	1
2. Nicotinic		14
Skeletal muscles	[illegible]	7
	[illegible]	10
	[illegible]	4
	[illegible]	1
3. Central nervous system		49
	excessive dreaming	33
	insomnia	29
	jitteriness, restlessness, increased tension, emotional lability, tremulousness	29
	nightmares, frequently with talking in sleep	17
	headache	8
	increased libido	6
	giddiness	6
	drowsiness	5
	paresthesias	3
	mental confusion	2
	visual hallucinations	1
	tremor	1
	pain in legs, sciatic distribution	1
4. Skin		
	urticaria (may have been sensitive to peanut oil)	2
5. No symptoms		6

other actions in addition to antiesterase effects, which have complicated this sort of progress. *Fifth*, it must be mentioned that the phosphorus compounds do not only inhibit cholinesterases but also some other enzymes, the so-called aliesterase, the proteolytic enzyme chymotrypsin, and the B-type esterase of mammalian sera (Mendel, Myers, *et al.*, 1953).

Our simple picture, therefore, of reversible specific inhibitors chemically related to acetylcholine can well be retained, provided we remember the two types of attachment to the enzyme, and allow for a combination of inhibitor with receptor which is at least so slowly reversible as to be, for practical purposes, permanent. There are, however, important complications: most inhibitors are pretty active against both enzymes, a few much more against one than the other; and cholinesterase is not the only type of enzyme inhibited.

There is one inhibitor of cholinesterase particularly active against pseudo-esterase, which comes in a category by itself, viz. triorthocresylphosphate (TOCP). This has been known for some years to produce a central nervous paralysis; it was later found to be an inhibitor of cholinesterase and later still found to be particularly active against pseudo-cholinesterase in the brain (Earl and Thompson, 1952). The interest in it arises from the fact that it produces central nervous system lesions of a kind also seen with two of the anticholinesterases, which will be discussed later.

Effects of Inhibition of Cholinesterase

It is fairly straightforward to interpret most of the effects seen in the terms of our original simple picture as being due to accumulation of acetylcholine, so that the results of inhibiting cholinesterases in the body provide a concise exercise in the anatomy of cholinergic nerves. The careful studies by Harvey and his colleagues on the effect of DFP on man, still provide one of the best descriptions of these effects. (Table 1.) The most prominent effects are the muscarinic ones. Of the nicotinic actions of acetylcholine only that on striated muscle reveals itself at all obviously. Remaining is a group of central nervous system actions, some of a quite remarkable kind.

There has been a good deal of discussion as to whether the

brain or red-cell true cholinesterase (Frawley *et al.*, 1952). But it needs only a little reflection to see how hard it is to test such a correlation fairly: one needs to know the degree of inhibition of the cholinesterase at the particular site of interest, for it is known that it may vary considerably from tissue to tissue. The most detailed attempts, by Riker and Wescoe (1949) on the salivary gland, and by Koelle and colleagues (Koelle *et al.*, 1950; Kamijo and Koelle, 1952) on the ganglion and small intestine, have been more encouraging. It is disappointing that we cannot judge cholinesterase concentrations in the recesses of the brain or the muscles from assays on readily accessible serum or red cells, but it was rather optimistic to expect it.

On the whole, then, one can retain the belief that most of the actions are due to acetylcholine accumulation, and look to further work to establish the parallelism of inhibition and effect.

There is one result of giving cholinesterase inhibitors of a different kind, with some potential clinical importance. Serum cholinesterase can hydrolyse, quite rapidly, two other substances, succinylcholine, and local anaesthetics, especially procaine. After the administration of an inhibitor, it would be quite wrong to expect succinylcholine still to be transient in action, or to suppose that it will be possible to inject such large amounts of procaine into tissues (for local anaesthetic purposes) as one can normally do without producing systemic effects.

Demyelination by Anticholinesterases

It was, as described above, at first believed that chronic administration of anticholinesterases produced (except for a hypoplasia of spleen and thymus) no morphological effects other than those due to the usual accumulation of acetylcholine. This is still, generally speaking, true. But as a result of exposure to certain insecticides a number of cases of curious paralysis have been described which have led to the discovery that at least two of the newer anticholinesterases are able to produce, in animals, a demyelination of parts of the nervous system. These are DFP and Mipafox. The lesion is the same as was seen some years ago when TOCP was drunk in a ginger wine in the States;

phosphorus inhibitors exert a direct action, as well as causing acetylcholine to accumulate. Douglas and Paton (1954) went into this in some detail with TEPP so far as the neuromuscular junction is concerned, and failed to find any evidence for this; but some interesting points in the progression of TEPP poisoning emerged. Suppose you record the electric potential along a muscle fibre; normally it is the same all the way along. But if you give a drug such as decamethonium or acetylcholine, then the endplate regions become depolarized (i.e. negative to the rest of the fibre).

If TEPP is given, there develops a similar depolarization, but it is subject to an irregular waxing and waning depending on the amount of activity in the motor nerve. We know from the analysis of the effects of such depolarization that they will cause muscle fibres at first to contract, and later become inexcitable if it persists: so that you will get with such a fluctuation the sort of mixture of fasciculation and neuromuscular block which you see clinically. With a bigger dose a much larger, more prompt, and more persistent effect is seen. Repetitive stimulation of a given motor nerve produces a big increase in the depolarization: all these effects can be attributed to preservation of acetylcholine released at motor-nerve endings. Now, if the muscle is denervated, the depolarization produced by TEPP is not completely abolished. But this is not due to a direct action by TEPP, since if you now take blood from the animal, there is acetylcholine in it, enough to produce the effect recorded, and running a similar time course. In poisoning by TEPP, therefore, one's muscles are paralysed first by the acetylcholine released by the activity in one's own motor nerves, and secondly (with much deeper poisoning) by acetylcholine accumulating in the blood (from the portal system and elsewhere).

Although qualitatively one can account for the actions of inhibitors, there is often a failure to relate symptomatology quantitatively to the degree of cholinesterase inhibition. One can find abundant evidence, for instance, that the serum cholinesterases can be almost completely inhibited without producing symptoms. Again, in animals, there is no, or only a poor, correlation between intensity of symptoms and the reduction of

and tightness in the chest and ptosis (rather than the abdominal symptoms of the first case). This progressed to severe weakness and hypotonia but was successfully treated in the acute stage with atropine, the patient being discharged about a fortnight later. About a week after discharge, however, he noticed cramp-like pains in the calves and feet when riding a bicycle and began to get rather weak in his legs. In hospital he was found to have a partial motor paralysis with tender muscles but was electromyographically fairly normal. With lapse of time some recovery took place but he remained easily tired for a considerable time.

In animals it is known that a good many other potent anticholinesterases cannot produce this lesion. Nevertheless the fact that three agents, TOCP, Mipafox and DFP, of this group can do so, means that it has to be taken into serious consideration as a possible result of any cholinesterase action. Two rather different questions arise here. Taking the experimental demyelination first, *is it due to cholinesterase inhibition?*

The argument for it is that all the agents producing it are cholinesterase inhibitors; and indeed, because TOCP is particularly an inhibitor of pseudo-cholinesterase in the central nervous system it has been suggested that the demyelination is particularly due to interference with this enzyme, supposing it to have some special relationship to the nutrition of nervous tissue. But against this supposition are three points: (*a*) that there are at least six other anticholinesterases which do not induce demyelination; (*b*) that there is no parallel between the intensity and course in time of the effect of the inhibitor on the tissue cholinesterases and the actual appearance of the demyelination; (*c*) TOCP, although it is an anticholinesterase, does not, when given in doses which produce this paralysis, elicit any of the usual cholinergic symptoms. One is left suspecting therefore that the demyelination is not due to cholinesterase inhibition but rather to the inhibition of some other enzyme by these particular compounds. Remembering that they are not entirely specific, this is not unlikely

Secondly, in the clinical cases, *is the motor paralysis due to inhibition of cholinesterase?* It seems reasonably clear that the initial weakness could be attributed to it, since it was known that the

and the analogy became still more striking when it was found that TOCP was a powerful inhibitor of pseudo-cholinesterase, particularly of the white matter of the brain. It is seen particularly clearly in hens in which lesions, chiefly of the spino-cerebellar tract but also of other fibre groups and of sciatic nerve, have been produced; demyelination of the sciatic nerves of rats and rabbits has also been shown. This lesion does not come on for something like one or two weeks; so that in the hen after DFP there is a characteristic progression from the typical poisoning in a hen due to an acetylcholine-like action, with paralysis and stiffly extended limbs, moving on later to ataxia and an incomplete motor palsy quite unlike the initial picture (Barnes and Denz, 1952).

The human cases after Mipafox followed a course rather like the hen (Bidstrup, Bonnell and Beckett, 1953). After the exposure to the insecticide the subjects complained of vomiting and diarrhoea and nausea; their pupils were tightly constricted and conjunctivae flushed; their muscles, particularly in the face and neck, were twitching and the whole musculature was weak. They were treated with atropine, 1/50th of a grain roughly every two hours which relieved the muscarinic symptoms but did not influence the muscular weakness and general hypotonia. Then about 15 days after the original exposure an ataxia was noticed. This slowly developed, producing a considerable weakness in the legs and a development of weakness in the hands and arms. Clinically there was a flaccid paralysis of the legs, the muscles were tender, there were twitchings in many muscles. Sensation was normal, tone considerably reduced. At this point there were no fibrillation potentials and the electromyographer thought the condition was compatible with a depolarization neuromuscular block. A particularly interesting point was that an attempt at physiotherapy, the usual treatment adopted for such cases, led to a very severe deterioration with increased paralysis, generalized twitchings and clonic spasms. Later there was an obvious atrophy of some of the muscles, particularly small muscles of the hand, but eventually recovery began and has progressed continuously.

A second case had a slightly different onset with wheezing

central nervous system. With OMPA, in a proportion of cases at least, a smooth and effective control can sometimes be established. Whether the extra trouble in controlling the dosage of a drug with such prolonged action and consequent danger of accumulation is worth while, when successful therapy with neostigmine can be achieved fairly easily, is uncertain. But what I wish to mention here is a very instructive instance of how confusing neuromuscular paralysis may be.

This can be illustrated with a case from America (Wilson *et al.*, 1952) of a man not responding very well to neostigmine, so that it was decided to try him on OMPA. He was started on 7 mgm. OMPA twice a day orally together with a neostigmine supplement, 1 mgm. approximately 4-hourly intramuscularly. In about a week his serum cholinesterase was down to 19 per cent of normal and red-cell cholinesterase to 10 per cent of normal. The control was still not adequate and the OMPA was increased to 14 mgm. twice a day, atropine was given to control salivation and neostigmine was slowly stopped. So far all was well. Then the patient began to get weaker again and neostigmine was given promptly intramuscularly to overcome what was thought *might* be a 'myasthenic crisis'. A few hours later the patient was weaker still, and the doctor in charge, now convinced that this was a true 'myasthenic crisis', gave yet more neostigmine intramuscularly. The patient now became weaker still, unable to swallow the saliva which was pouring from his mouth faster than it could be sucked out and sweating so hard that the bed linen was wringing wet. Fortunately, a consultation now took place and the diagnosis was made of a 'cholinergic crisis'. Administration of anticholinesterases was immediately stopped; 1/50th of a grain of atropine was given intravenously; and the patient was put in an oxygen tent. The atropine was repeated every 4 hours and in 36 hours he had greatly improved. Eventually he was returned to a maintenance dose of 10 mgm. of OMPA night and morning.

There are three interesting points here. *First*, that the onset of poisoning by an anticholinesterase can occur in a myasthenic patient, and appears superficially as an intensification of the myasthenic process. *Secondly*, it was noted that there was no use-

blood cholinesterase was reduced, that there were often fasciculations and cramps, and that the condition worsened with muscular activity. But it is difficult to see how cholinesterase inhibition by itself, at the motor endplates, can account for a sequence of events whereby the initial muscular weakness virtually passes off and is then re-established again accompanied by wasting of muscles and ultimately muscular fibrillations, without any signs of cholinesterase inhibition anywhere else in the body. The whole picture suggests, in short, that for the first week or two there is a true antiesterase neuromuscular block but that this slowly moves over into a paralysis due to a partial denervation possibly produced by a destruction of nerve trunks remote from the muscle.

Whatever may be the truth of these speculations, two practical points arise: *first*, that every case of anticholinesterase poisoning should be followed for one or two months after the acute stage has been resolved; and *second* that if a paralysis does develop, it should be treated rather cautiously and without the vigorous physiotherapy so useful in other lower motor neurone lesions.

Treatment of Myasthenia Gravis

The basis of therapy of myasthenia gravis by antiesterases is fairly well agreed. Until recently neostigmine was virtually the only drug of importance, being much superior to eserine because of its greater proportionate effects on muscles. Its derivative—tensilon—has not really supplanted it, despite some attractive features, because it tends to be too transient in action. A weakness of neostigmine therapy, however, is that the patient is condemned to continual ups and downs in his muscular strength, because of the fairly short duration of action of the drug. When the irreversible inhibitors appeared, therefore, it was hoped that these might allow a smoother, more continuous control. DFP, the first one used, was reasonably effective and gave a more sustained action than neostigmine; but it did not produce as good a result as neostigmine at the peak of its action, and in addition exerted the somewhat unpleasant central effects previously mentioned. More promising has been the use of OMPA which, because of its low fat solubility, does not penetrate the

and respiratory effects rather than to the muscarinic actions, and so far as broncho-spasm or fall in blood pressure contribute to the lethal process, it may not be much reduced.

Antagonists to the Effect of Cholinesterase Inhibitors

Cases of poisoning by anticholinesterases may occur during the use of insecticides in agriculture, during treatment of myasthenia, during surgical anaesthesia and (perhaps) in wartime, so that the importance of treatment of such poisoning needs no emphasis. There is some debate whether death, when it occurs, is due to central respiratory depression, peripheral neuromuscular block, bronchospasm, circulatory failure, or pulmonary oedema, or to an assortment of these. But this is a somewhat academic debate, since cases vary so much that one must clearly be ready for all the important manifestations of the poisoning.

The pre-eminent remedy is atropine, principally for the muscarinic actions (cf. Grob, 1950). It has repeatedly proved itself able to relieve or abolish these and relatively large doses can be given. The usual advice is to give 1/50th of a grain intravenously if necessary, perhaps every 15 minutes if poisoning is severe, until there are clear signs of atropinization, such as dilation of the pupil. Atropine has little effect on the neuromuscular block but it is claimed that it can relieve the depressant effect of anticholinesterases on the central nervous system (Douglas and Matthews, 1952). Whether this is important clinically is not established, but it is certainly not an undesirable effect.

The treatment of neuromuscular paralysis is distinctly harder. It is known that d-tubocurarine can prevent many of the effects of inhibitors at the neuromuscular junction; but for practical purposes the obvious difficulty occurs that it itself produces neuromuscular block, so that although the inhibitor has been antagonized, the patient is little better off. Parkes and Sacra (1954) have shown that hexamethonium can increase the protection given by atropine (normally about two-fold) to as much as ten-fold, and that it confers considerable protection by itself (this may well be due to its feeble neuromuscular action which forms the basis for its original suggested antagonism to decamethonium). The protection by atropine and hexamethonium

ful warning change in the cholinesterases values of serum and blood cell. These remained at more or less the same rather low levels both before and during the crisis. *Third*, atropine appears to have been an excellent antidote and fully justified its universal recommendation. There are indications that similar crises occur when over-enthusiastic attempts are made to antagonize the effects of curarizing compounds given during surgical anaesthesia.

Accommodation to Anticholinesterases

It must not be assumed that the body accepts the insults of anticholinesterase administration completely passively. There are two interesting bits of evidence pointing to the contrary. First, it has been found that after the administration of a number of these inhibitors, the depression of cholinesterase activity may be succeeded by a rise in cholinesterase activity of a tissue. The conditions governing this seem still somewhat obscure but it is [illegible] that a [illegible] cur in [illegible] 1952).

A better established observation with a number of inhibitors is that, if an animal can be tided over the first exposure, he may become relatively resistant to a subsequent one. For instance, Barnes has shown (1953) that if artificial respiration was given after a normally lethal dose of paraoxone, recovery takes place in approximately 25 minutes. If now the dose is repeated half an hour or an hour later, there is less muscular fasciculation and collapse and the depression of the breathing is slower and less intense. Corresponding to this is an observation made by Douglas and Paton (1954) that the depolarization of the motor endplate produced by large doses of TEPP tends to wane with successive administrations. This is not because the amount of acetylcholine accumulating in the blood changes, for one can find just as high levels after subsequent doses as after the first one. There seems to be some real accommodation by the endplate to acetylcholine of the same sort as Burns and Paton (1951) observed with decamethonium at the endplate. This accommodation seems to be restricted to the neuromuscular

and respiratory effects rather than to the muscarinic actions, and so far as broncho-spasm or fall in blood pressure contribute to the lethal process, it may not be much reduced.

ANTAGONISTS TO THE EFFECT OF CHOLINESTERASE INHIBITORS

Cases of poisoning by anticholinesterases may occur during the use of insecticides in agriculture, during treatment of myasthenia, during surgical anaesthesia and (perhaps) in wartime, so that the importance of treatment of such poisoning needs no emphasis. There is some debate whether death, when it occurs, is due to central respiratory depression, peripheral neuromuscular block, bronchospasm, circulatory failure, or pulmonary oedema, or to an assortment of these. But this is a somewhat academic debate, since cases vary so much that one must clearly be ready for all the important manifestations of the poisoning.

The pre-eminent remedy is atropine, principally for the muscarinic actions (cf. Grob, 1950). It has repeatedly proved itself able to relieve or abolish these and relatively large doses can be given. The usual advice is to give 1/50th of a grain intravenously if necessary, perhaps every 15 minutes if poisoning is severe, until there are clear signs of atropinization, such as dilation of the pupil. Atropine has little effect on the neuromuscular block but it is claimed that it can relieve the depressant effect of anticholinesterases on the central nervous system (Douglas and Matthews, 1952). Whether this is important clinically is not established, but it is certainly not an undesirable effect.

The treatment of neuromuscular paralysis is distinctly harder. It is known that d-tubocurarine can prevent many of the effects of inhibitors at the neuromuscular junction; but for practical purposes the obvious difficulty occurs that it itself produces neuromuscular block, so that although the inhibitor has been antagonized, the patient is little better off. Parkes and Sacra (1954) have shown that hexamethonium can increase the protection given by atropine (normally about two-fold) to as much as ten-fold, and that it confers considerable protection by itself (this may well be due to its feeble neuromuscular action which forms the basis for its original suggested antagonism to decamethonium). The protection by atropine and hexamethonium

together can thus be a notable one, although it is less striking when given *therapeutically*.

A number of other compounds have been suggested as antagonists. Nikethamide, for instance, although itself useless as a protective agent, when given to rats together with atropine (on the theory that it would diminish the central depressant effects) was found to increase atropine's protective effect by eight-fold in female rats and two-fold in male rats (DiStefano *et al.*, 1951). Similarly Heymans (1950) has claimed that Diparcol and some of the other nicotinolytic drugs can confer a useful protection. Other analeptics, such as metrazol, benzedrine, picrotoxin or caffeine, are all believed to increase mortality.

The general conclusion is that atropine is outstanding as a therapeutic agent, that hexamethonium might be tried cautiously where there is confidence that neuromuscular block is an important element in the poisoning, but that the use of other agents should be deferred until more work has been done. It is, of course, a truism, but one worth repeating, that other measures in the treatment of a patient should not be forgotten: (1) Artificial respiration with oxygen to counter the effects of respiratory paralysis, bronchospasm (cf. Holmstedt, 1951), or pulmonary oedema; (2) appropriate measures for raising the blood pressure if this sinks very low and cannot be controlled by atropine (although this seems very rare); and (3) care that the patient is not getting cold; this may easily happen with profuse sweating, and perhaps some peripheral cutaneous vasodilation.

One may conclude, perhaps, by summarizing the possible ways of protection available at present against such 'nerve gases' in this way. Let us suppose that you are a sufferer from Parkinson's disease (therefore taking diparcol), a severe hypertensive (therefore on hexamethonium), a myasthenic (therefore taking neostigmine and atropine), and with cirrhosis of the liver (and therefore unable to activate Parathion or OMPA); then you may well feel proud, for if there is another war, your extraordinary resistance to these war poisons will ensure your early despatch not to hospital, but to the 1st Commando regiment.

REFERENCES

Aldridge, W. N. and Davison, A. N. (1952). *Biochem. J.* **52**, 663–71.
Augustinsson, K. B. (1948). *Acta physiol. scand.* **15**, Suppl. 52.
Austin, L. and Berry, W. K. (1953). *Biochem. J.* **54**, 695–700.
Barnes, J. M. (1953). *Brit. J. Pharmacol.* **8**, 208–11.
Barnes, J. M. and Denz, F. A. (1952). *J. Path. Bact.* **65**, 597–605.
Bidstrup, P. L., Bonnell, J. A. and Beckett, A. G. (1953). *Brit. med. J.* **i**, 1068.
Burgen, A. S. V. and Chipman, L. M. (1952). *Quart J. exp. Physiol.* **37**, 61–74.
... *Physiol* **115**, 41–73
... 19
... and Hodge, H. C. (1951). *Proc. Soc. exp. Biol, N.Y.* **78**, 712–3
...
105, 156–65.
Funke, A., Depierre, F. and Krucker, W. (1952) *C.R. Acad Sci., Paris,* **234**, 762–4.
Gardiner, J E. and Kilby, B. A (1952). *Biochem. J.* **51**, 78–85.
Grob, D., Lilienthal, J. L., Harvey, A. M. and Jones, B F (1947) *Bull. Johns Hopk. Hosp* **81**, 217–44.
Grob, D. *et al.* (1950). *J Amer Ass* **144**, 104–8.
Harvey, A. M., Lilienthal, J L., Grob, D., Jones, B. F. and Talbot, S. A. (1947) *Bull. Johns Hopk Hosp.* **81**, 267–92
...
–57.
... *J Pharmacol.* **100**, 180–91
Locker, A and Siedeck, H. (1952). *Experimentia*, **8**, 146–8.
Mendel, B, Myers, D K., Uyldert, I. E, Ruys, A. C and de Bruyn, ...
... *acol.* **9**, 299–305.
... –61.
... *J Pharmacol.* **4**, 381–400.

Pelikan, E. W., Smith, C. M. and Unna, K. R. (1954). *J Pharmacol.* **111**, 30–40.
Riker, W. and Wescoe, W. C. (1949). *J Pharmacol.* **95**, 515–27.
Whittaker, V. P. (1951). *Physiol. Rev.* **31**, 312–43.
Wilson, C. W., Williams, J. P. and Miller, D. H. (1952). *Ann. intern. Med.* **37**, 574–9.

VIII

Acetylcholine and the Maintenance of the Cardiac Rhythm

J. H. BURN

THE general hypothesis which is to be discussed is that the contractions of the heart are initiated by the local formation of acetylcholine, and that acetylcholine is concerned in the conduction of the impulse from the pacemaker to the ventricles. In skeletal muscle we know that the nerve impulse is transmitted by the liberation of acetylcholine from the nerve-ending; the hypothesis suggests that in cardiac muscle acetylcholine plays a similar role but that the acetylcholine is formed locally.

In 1946 Dawes published a paper on substitutes for quinidine in which he pointed out that quinidine is a substance which antagonizes the action of acetylcholine in all forms of muscle. Thus in the presence of quinidine the stimulant action of acetylcholine on the frog's rectus abdominis and on the isolated intestine of the rabbit is reduced; similarly, in the presence of quinidine the inhibitory action of acetylcholine on the auricles of the rabbit heart is reduced. These observations led to the idea that perhaps the use of quinidine to suppress auricular fibrillation depended on the power of quinidine to antagonize acetylcholine, and that despite all appearance to the contrary, acetylcholine was responsible for the maintenance of cardiac contractions. The appearances are indeed very much against such a view, for the ordinary effect of acetylcholine on heart muscle is to inhibit it; to suggest that acetylcholine maintains the cardiac contractions is like suggesting that the brake on the wheel is the mechanism which drives the motor-car.

PELIKAN, E. W., SMITH, C. M. and UNNA, K. R. (1954). *J. Pharmacol.* **111**, 30–40.
RIKER,
WHITTA
WILSON
37, 574–9.

become feebler and then stop. The addition of acetylcholine to the bath then causes the beat to begin again. Further additions of acetylcholine augment the rate and amplitude of the contractions, but a point is reached when the inhibitory effect returns. This stimulant action of acetylcholine was originally observed in fourteen out of twenty-two preparations, and has been often confirmed since.

The demonstration of the stimulant action when the auricles had ceased to beat suggested that the choline acetylase mechanism was then failing, and that contractions were absent because insufficient acetylcholine was synthesized to initiate a propagated disturbance in the heart muscle. We tested this suggestion by determining the choline acetylase activity (1) of a series of freshly excised auricles, (2) of a series of auricles which had ceased to beat after being 24–30 hours in the bath, and (3) of a series of auricles which, having stopped, had then resumed their contractions on the addition of acetylcholine. We found (Bulbring and Burn, 1949) that the mean choline acetylase activity in the fresh auricles was 46 μg./g./hr., in the stopped auricles was 9 μg./g /hr. and in the restarted auricles was 37 μg./g./hr. To observe that choline acetylase activity was very low in the auricles which had ceased to beat was not surprising, but to find that when the beat began again the choline acetylase activity was soon restored to a figure approaching normal was to obtain striking support for the view that the maintenance of the heart beat was closely connected with, and perhaps dependent on, the synthesis of acetylcholine.

A further correlation between the activity of the auricles and acetylcholine synthesis was obtained when we studied the effect of acetylcholine itself. When added to the bath, acetylcholine causes a decline in the rate and amplitude of the beat; when added to the tube in which the auricle powder was incubated with choline, acetylcholine was found to depress the rate of synthesis

Action of Eserine on the Heart Rate

Evidence was thus obtained that the choline acetylase system was related to cardiac activity, and we then wished to discover if the same was true for the cholinesterase system. For if this

The experiments to be described have been carried out for the most part on the isolated auricles of the rabbit heart. These when dissected from the freshly excised heart were suspended in a well-oxygenated bath of Locke's solution at 29°C. Under these conditions they maintain spontaneous contractions for as long as thirty hours.

Acetylcholine is present in the auricles. This can be verified by grinding them up in acid saline containing eserine, or in acid alcohol. When suitably treated the extract can be shown to contain an acetylcholine-like substance by biological tests.

Acetylcholine is formed in the body by an enzyme choline acetylase which can attach an acetyl group to choline, choline being itself an almost inactive substance. Choline acetylase is present in the auricles as was shown by Comline (1946), and its activity can be measured by the method described by Feldberg and Mann (1946). The method consists in extracting the ground-up auricles with acetone, and using the dry powder left behind for incubation together with choline, adenosinetriphosphate (to supply energy) and sodium citrate (to supply acetyl groups). The enzyme present in the powder then synthesizes an amount of acetylcholine which is measured. The choline acetylase activity of the powder prepared from the auricles can then be expressed as the number of microgrammes of acetylcholine synthesized per gramme of powder per hour.

Acetylcholine is destroyed in the body by cholinesterase, and its presence in the auricles of the rat was shown by Ord and Thompson (1950). Cholinesterase is present in the auricles of the rabbit in both of the two forms 'true' and 'pseudo'. The hypothesis that the heart beat is maintained by acetylcholine could not be maintained if acetylcholine were not present, or if either of the enzymes choline acetylase and cholinesterase were absent; in fact all three are there.

Stimulant Action of Acetylcholine

Under certain circumstances the usual inhibitory action of acetylcholine on the spontaneously beating auricles is reversed. My colleague, Dr. Edith Bulbring, observed that when the auricles are left in the bath for 24–30 hours, the contractions

Eserine was observed to cause slowing of the rate as the mean effect in several auricles when it was allowed to act for 30 minutes in concentrations ranging from 10^{-7} g./ml. to 10^{-4} g./ml., the slowing of the rate being progressively greater as the concentration of eserine was increased. When the concentration was raised further to 10^{-3} g./ml., after 5–15 minutes the contractions stopped abruptly. On the amplitude of the contractions the effect was different, since it was found that concentrations of eserine up to 10^{-4} g./ml. caused the amplitude to increase. Above this concentration, however, the amplitude was reduced. A similar effect on the heart rate, though not on the amplitude, was observed when di-*iso*propylfluorophosphonate (D.F.P.) was used instead of eserine. In a certain proportion of trials when eserine was tested on the auricles, the rate was accelerated. This occurred in 10 out of 45 trials, when the concentrations of eserine varied from 10^{-7} to 10^{-5} g./ml. When a concentration such as 10^{-6} was found to increase the rate, a higher concentration reduced it.

Action of Eserine on the Frog Heart

A similar variation occurred in the action of eserine in concentrations 10^{-7} g /ml and 10^{-6} g./ml. when perfused through the isolated frog heart. The usual effect, however, was an increase in the heart rate seen in the first 5–10 minutes of the perfusion. With continued perfusion the rate fell, sometimes to a point below the initial rate, with a return to the initial rate when the perfusion with eserine was stopped at the end of 30 minutes. In some experiments the effect of eserine was to cause a rise in the rate throughout the period of perfusion, and in others to cause a fall.

These different effects were significant. The fact that eserine slowed the rate in the heart-lung preparation indicated that acetylcholine was being formed in the heart, but it did not indicate at what point. The point could have been the pacemaker itself, or it could have been at a point outside the pacemaker. If the pacemaker itself initiated the beat without the aid of the hypothetical acetylcholine mechanism, the effect of acetylcholine arriving at the pacemaker would be

system was responsible for destroying the acetylcholine formed, it should be possible to modify cardiac function by inhibiting the cholinesterase. A study was therefore made by Dr. J. M. Walker and myself in the heart-lung preparation of the dog prepared as described by Starling. Since the central nervous system is excluded from the circulation, the heart rate and the force are independent of external control; in addition the vagi are cut. The heart pumps the blood through an artificial resistance, which can be set at will at any height, into a reservoir from which it returns to the superior vena cava.

In a series of experiments in which the preparation was observed from 20–30 minutes to ensure that the heart rate was steady, eserine sulphate was added to the blood in a concentration of 10^{-6} g./ml. The heart rate then fell in the course of the next 5–10 minutes to a mean figure of 62 per cent of the initial rate. The heart continued to beat at the slower rate, until a small amount, 10 µg., of atropine sulphate was injected into the blood going to the heart, when the rate rose rapidly to its previous value. A similar result was obtained when other anticholinesterases were used, such as neostigmine. The slowing of the rate by the eserine was taken to indicate that acetylcholine was being steadily produced in the heart, and that when its destruction was delayed, the heart rate fell in consequence. It was conceivable that this acetylcholine might come from the vagus nerve-endings in the form of a slight leak continuing after section of the vagi. This point was tested in a series of preparations in which the vagi were cut four days earlier; in these preparations eserine caused a similar fall in rate. We were further able to show that in preparations in which hexamethonium was injected so that the strongest stimulation of the peripheral end of the vagus had no effect on the heart rate, the injection of eserine still caused the usual fall. We therefore concluded that the formation of acetylcholine taking place in the heart was not due to the extrinsic nerves.

Action of Eserine on Auricles

Observations of the action of anticholinesterases were also made on the isolated rabbit auricles (Burn and Kottegoda, 1953).

beat stopped. When acetylcholine was added to the bath (the quinidine being still present) the beats began again. A variant of this was the addition of acetylcholine when, under the influence of quinidine, the rate of beating had become very slow. The immediate effect of the acetylcholine was to cause a further slowing of the rate, but this was transient and both rate and amplitude soon became greatly accelerated. The results could be explained by the hypothesis that the contractions were initiated by the formation of acetylcholine within the tissue of the auricles, and that quinidine steadily reduced the action of this intrinsic acetylcholine by occupying the receptors on which it acted. When the action of quinidine had proceeded to arrest of the contractions, these began again as soon as the acetylcholine concentration was raised by the addition of more to the bath. It is important to point out that when the auricles had not completely stopped, the usual inhibitory action of acetylcholine was observed as a transient effect before the stimulant action on rate and amplitude was seen.

Resemblance of Quinidine to Eserine

The findings with quinidine led to the idea that the arrest of the auricles by a high concentration of eserine which was described above might have a similar explanation. For if quinidine caused arrest of the auricles by occupying receptors on which acetylcholine acted, then other substances which possessed an affinity for these receptors would have the same action. Since cholinesterase combines with both acetylcholine and with eserine, eserine would be expected to occupy receptors on which acetylcholine acts We therefore exposed auricles to a high concentration of eserine until the contractions stopped, and we then added acetylcholine to see if the contractions would start again. We found that provided the concentration of acetylcholine added was not too great, the contractions were resumed. This result appeared to us to provide considerable support for the general hypothesis put forward, for the result did not seem otherwise explicable.

Further evidence that the action of eserine in high concentration resembled that of quinidine was provided by the observa-

expected always to be a slowing of the rate, as was observed. Now in the isolated auricles, beating in a bath of Locke's solution at 29°C., and in the frog heart perfused with Ringer's solution, eserine caused not only slowing of the rate, but also acceleration. Such an acceleration could not have been produced by acetylcholine arriving at the pacemaker from a point external to it, but it could have been produced if the acetylcholine were formed at the pacemaker. It is conceivable that the conditions provided by the bath for the auricles and by the perfusion fluid for the frog heart were sometimes suboptimal, and that the formation of acetylcholine was less rapid than in an environment of blood. In such circumstances the presence of eserine, by diminishing the rate of destruction of acetylcholine, would increase the rate of beating, since it would shorten the time necessary for a threshold concentration of acetylcholine to be formed.

The Action of Quinidine

A series of observations was then made to investigate the action of quinidine on the auricles. It has already been pointed out that quinidine reduces the action of acetylcholine in all forms of muscle, from which it may be concluded that quinidine competes for the receptors on which acetylcholine acts. When a small amount of acetylcholine is added to the bath containing the auricles, such that the final concentration is $2 \cdot 5 \times 10^{-8}$ g./ml., there is a slight inhibition of the amplitude of the contractions. If, however, the same amount of acetylcholine is added in the presence of eserine 10^{-6} g./ml. the beats are arrested. This observation was made by Webb (1950). We found that when quinidine was added to the bath at this point, in concentration 10^{-5} g./ml., the beats began again. Before the quinidine was added, the amount of acetylcholine present was sufficient to maintain arrest of the beat; when quinidine was added, the effect of this acetylcholine was reduced, and it was no longer able to maintain arrest of the beat, which therefore began again.

A further observation (Briscoe and Burn, 1953) was that when the actively beating auricles were exposed to quinidine (10^{-5} g./ml.) their rate and amplitude gradually diminished until the

the injection of 4 μg. acetylcholine. In other experiments in the presence of eserine 10^{-6}, the injection of 1 μg. and of 2 μg. acetylcholine caused a regular beat to become temporarily irregular.

SUMMARY

The results described all point in favour of the hypothesis that the local formation of acetylcholine is responsible for the contractions of the heart. There is reason to suppose that acetylcholine is formed at the pacemaker, and that when a certain concentration is reached a contraction wave is propagated through the auricles. There is also reason to believe that acetylcholine is concerned in the conduction of the wave. Evidence in support of these conclusions comes from the observation that auricles which have ceased to beat will start again if acetylcholine is added. Also from the observation that the activity of the choline acetylase system is related to the activity of the auricles. Thus the choline acetylase system is active in fresh auricles; is inactive in stopped auricles; is once more active in auricles restarted by acetylcholine. Evidence also comes from [illegible]

that of high concentrations of eserine in the isolated auricles. Evidence finally comes from the effect of eserine in converting an irregular beat to a regular beat in the isolated auricles, and to the action of acetylcholine in relation to irregularity in the heart-lung preparation. Much more work still lies ahead, for there are many unanswered questions Hitherto, however, there have been no results which conflict with the hypothesis put forward.

REFERENCES

BRISCOE, S and BURN, J H (1950). *Brit J Pharmacol.* **9**, 42
BÜLBRING, E and BURN, J H (1949) *J Physiol* **108**, 508.
BURN, J H and KOTTEGODA, S R (1953) *J. Physiol.* **121**, 360.

tion that the action of the two substances was additive. In the presence of a concentration of eserine *too low to cause arrest of* the auricles, the addition of quinidine stopped the auricles within a minute or two, instead of taking 20–30 minutes to do so.

Failure of Conduction as Cause of the Arrest

To explain the effects of quinidine and of high concentrations of eserine in causing arrest of the auricles is difficult without analysis of the electrical changes which take place. Quinidine has long been known to delay conduction, and the arrest of the beat may be due to failure of conduction. We observed that auricles which were arrested by eserine did not respond to electrical stimulation, that is to say, an electrical stimulus was not conducted, and eserine has been shown by Rothschuh and Bammer (1952) to cause delay in conduction in strips of frog ventricle. These *authors also found that neostigmine increased the rate* of conduction in the frog ventricle, and we have observed that even the highest concentrations of neostigmine did not cause arrest of the auricles. All these points unite to support the view that quinidine and eserine arrest the auricles by causing failure of conduction, and acetylcholine since it improves conduction may cause the beats to be resumed by doing so.

Effect of Eserine on Irregularity

If the heart beat is due to the local formation of acetylcholine, irregularities of the beat might be expected to arise from insufficient or irregular formation of acetylcholine. We have observed that the auricles when suspended in Locke's solution sometimes beat irregularly, and that the beat became regular when eserine (e.g. in concentration 10^{-6} g./ml.) was added. The regularity was in some experiments accompanied by a slowing of the rate, and in others by a quickening of the rate, the latter being observed when the irregular rate was very slow.

In the heart-lung preparation we have observed that acetylcholine in certain cases corrected irregularity, but in others, when eserine was present, caused irregularity. In one experiment a persistent 2 : 1 block was converted to normal rhythm by

IX

The Growth Hormone of the Anterior Pituitary Gland

F. G. YOUNG

History

THE genesis of our present knowledge of the pituitary growth hormone may be ascribed to the recognition, by Pierre Marie in 1886, of the syndrome of acromegaly. In 1887 the association of acromegaly with changes in the pituitary fossa was recorded by Minkowski while already, in 1884, the observation by M. Loeb, subsequently confirmed by Marie, Kanthack and others, that diabetes is common in patients with pituitary tumours presented the first indication of the important role in carbohydrate metabolism which the pituitary gland is now known to enjoy. At present it is widely accepted that gigantism and its opposite, dwarfism, often result from dysfunctions of the anterior pituitary gland, although uncomplicated oversecretion or undersecretion of the anterior pituitary growth hormone are conditions probably only seldom encountered in the human being.

The modern era of research on the anterior pituitary gland may be dated from the perfection by P. E. Smith in the early 1920's of a method of surgical hypophysectomy in the rat. In his classical studies Philip Smith described the alterations in many other endocrine glands which follow removal of the pituitary gland, and showed how these could be prevented or cured by the implantation of rat anterior pituitary tissue into the pituitaryless rat. The most striking changes of this sort were found in the thyroid gland, in the adrenal cortex and in the

COMLINE, R. S. (1946). *J. Physiol.* **105**, 6P.
DAWES, G. S. (1946). *Brit. J. Pharmacol.* **1**, 90.
FELDBERG, W. and MANN, T. (1946). *J. Physiol.* **104**, 411.
ORD, M. G. and THOMPSON, R. H. S. (1950). *Biochem. J.* **46**, 346.
ROTHSCHUH, K. E. and BAMMER, H. (1952). *Z. ges. exp. Med.* **119**, 327.
WEBB, J. L. (1950). *Brit. J. Pharmacol* **5**, 335.

to the somewhat crude method ascribed to Archimedes. In an attempt to determine whereabouts in the body of the rat the protein was deposited under the influence of growth hormone, Dr. Greenbaum and I (Greenbaum and Young, 1950, 1953) studied the size and protein content of different muscles in the growth-hormone treated normal rat, and found that under the influence of growth hormone some muscles grew significantly faster than the body as a whole, while others tended to lag behind. The possibility occurred to us that the muscles which grew most readily under the influence of growth hormone possessed protein of greater lability than did other tissues, and we therefore studied the response to starvation for six days of the muscles of the normal rat. We found that, with the exception of the masseter muscle, those tissues which gained protein most rapidly under the influence of growth hormone were those which lost it most rapidly during starvation, and conversely. We concluded that our results showed the existence of a group of muscles with protein of greater lability than that of the rest of the body tissues, these muscles responding most vigorously to the deposition of protein during treatment with growth hormone, and suffering most depletion of protein during inanition. Later studies showed that the composition of such a muscle was, in the growth-hormone treated rat, not significantly different from normal (Gray and Young, 1954).

For the assay of growth hormone one does not, of course, normally analyse the whole body, but one uses simple criteria of growth such as the increase in weight of the whole body, either in a growth-hormone treated normal animal or in a hypophysectomized one. The widening of the epiphyseal line in the tibia of the hypophysectomized rat under the influence of growth hormone has also been utilized, particularly by H M. Evans and C. H. Li. A rise in the plasma inorganic phosphate level has been employed as an indicator of growth hormone action by some investigators particularly in clinical investigations (cf Reifenstein, Kinsell and Albright, 1946), although this is of course much more indirect than the other methods mentioned above

[illegible]

pancreas.

Previously, in 1922, H. M. Evans and J. A. Long had described the promotion of growth which could be induced by the intraperitoneal injection of a crude anterior pituitary extract into normal rats. This observation, together with those of Philip Smith, confirmed and extended what was already surmised from the early clinical findings and from the experimental work of Cushing, Aschner and others, namely that the anterior pituitary gland secretes a hormone which stimulates the growth of the body as a whole.

Growth and Body Composition

Growth is of course a complex function, but it is evident that the deposition of protein must be an essential element in cellular enlargement and mitosis. In 1936 Lee and Ayres demonstrated clearly that hypophysectomy in the rat results in a loss of protein to the body as a whole, with a gain in fat and a rise in the total amount of energy stored. The experiments of Lee and Ayres were carefully controlled, particularly with respect to food intake, because removal of the pituitary gland depresses the appetite of an animal. Since hypophysectomy diminishes the total amount of protein in the body it is not surprising to find that treatment with growth hormone, both in the normal and in the hypophysectomized animal, raises the protein content of the body (see Young, 1953a).

When normal rats are treated with growth hormone they may take on an obese appearance although the fat content of the body is in reality diminished. This diminution in fat content, with an increase in protein, causes a rise in the specific gravity of the carcass (Young, 1945), and it is surprising that measurements of the specific gravity have not been used more as an indicator of the relative protein and fat contents of the body. The undertaking of such measurements in clinical medicine might have some difficulties, but nevertheless the specific gravity of the human being can be measured without recourse

dition of diabetes. Subsequently, the administration of a suitable crude anterior pituitary extract to intact dogs was found to produce a diabetic condition which persisted indefinitely after cessation of the injection of the extract (Young, 1937). The ability of growth hormone to induce diabetes in some circumstances satisfactorily accounts for the recorded frequent occurrence of diabetes in acromegaly. Indeed, one may wonder why diabetes does not occur in every case of this disease. The failure of diabetes always to develop in acromegaly must, however, be considered in relationship to the relative sensitivities to hormonal influences of the enzyme systems in the tissues under different conditions, as well as with respect to the different degrees of hormonal activity that are met with in disease.

How is it that growth hormone can be both diabetes-inducing

which suggests that two or more physiological activities may be ascribed to a single protein hormone. The number of different physiological actions attributed to growth hormone has already reached double figures and is still increasing. It is not, however, universally accepted that all the effects ascribed at one time or another to it are indeed the properties of the hormone itself. In particular, Recant has obtained evidence which suggests that the R.Q. depressing action of some preparations of growth hormone is not due to the hormone itself, and it seems possible that the myoglycostatic and R.Q. depressing factors are identical. As Reid, Smith and Young (1948) first pointed out, the diabetogenic agent in crude alkaline ox pituitary extracts (now realized to be mainly if not entirely growth hormone) is not identical with the pituitary factor which, in a limited proportion of instances, is found to reversibly inhibit *in vitro* the activity of hexokinase. Nevertheless, a substantial number of different types of physiological activity are at present ascribed by responsible investigators to growth hormone, and we must inquire how far the evidence will allow us to attribute all these actions to the hormone itself with certainty.

The criteria of purity of proteins are notoriously unsatisfac-

The Nature of Growth Hormone

In 1945 Li, Evans and Simpson succeeded in isolating a crystalline protein from anterior pituitary tissue, which they believed to be pure growth hormone. Subsequently Wilhelmi and others obtained a similar protein by different methods, and it appears from the evidence at present available that this protein has a molecular weight of about 47,000 and an isoelectric point of 6·85. Since its first isolation a very large number of different physiological effects have been ascribed to this protein, some of which are given in Table 1.

Table 1. Physiological Actions Ascribed to Growth Hormone

1. Stimulation of growth
2. [illegible]
3.
4.
5.
6. Depression of glucose uptake by isolated diaphragm (after parenteral administration)
7. Preservation of muscle glycogen during fasting (myoglycostatic effect)
8. Depression of R.Q.
9. Stimulation of insulin secretion
10. [illegible]
11.
12.
13.
14.
15.

The Diabetogenic Action of Growth Hormone

The diabetogenic effect of growth hormone, first found by Cotes, Reid and Young (1949) on the administration of purified growth hormone to the intact cat, is one that has particularly interested us, since it appears that growth hormone is undoubtedly the chief, if not the only, diabetogenic agent responsible for most of the effects first described by Houssay and his collaborators from 1930 onwards. You will recall that Houssay and his colleagues showed that hypophysectomy ameliorates an existing pancreatic diabetes and that the administration of a crude anterior pituitary extract to partially depancreatized or even intact dogs will, under suitable conditions, produce a con-

dition of diabetes. Subsequently, the administration of a suitable crude anterior pituitary extract to intact dogs was found to produce a diabetic condition which persisted indefinitely after cessation of the injection of the extract (Young, 1937). The ability of growth hormone to induce diabetes in some circumstances satisfactorily accounts for the recorded frequent occurrence of diabetes in acromegaly. Indeed, one may wonder why diabetes does not occur in every case of this disease. The failure of diabetes always to develop in acromegaly must, however, be considered in relationship to the relative sensitivities to hormonal influences of the enzyme systems in the tissues under different conditions, as well as with respect to the different degrees of hormonal activity that are met with in disease.

How is it that growth hormone can be both diabetes-inducing

which suggests that two or more physiological activities may be ascribed to a single protein hormone. The number of different

another to it are indeed the properties of the hormone itself. In particular, Recant has obtained evidence which suggests that the R.Q. depressing action of some preparations of growth hormone is not due to the hormone itself, and it seems possible that the myoglycostatic and R.Q. depressing factors are identical. As Reid, Smith and Young (1948) first pointed out, the diabetogenic agent in crude alkaline ox pituitary extracts (now realized to be mainly if not entirely growth hormone) is not identical with the pituitary factor which, in a limited proportion of instances, is found to reversibly inhibit *in vitro* the activity of hexokinase. Nevertheless, a substantial number of different types of physiological activity are at present ascribed by responsible investigators to growth hormone, and we must inquire how far the evidence will allow us to attribute all these actions to the hormone itself with certainty.

The criteria of purity of proteins are notoriously unsatisfac-

tory, and if we must consider the possible presence of an impurity amounting to less than 5 per cent of the protein itself, great caution is needed. The recent history of the pituitary adrenocorticotropic hormone (ACTH) is a warning in this connection. The purified protein isolated by Li and his collaborators in 1945, and independently by Sayers and his colleagues in the same year, and called ACTH (protein hormone ACTH), obeyed all the criteria that were usually applied to the homogeneity of proteins. Nevertheless this so-called homogeneous protein has more recently been separated into two fractions, one more than a hundred times as active as the protein hormone, and the other consisting principally of protein almost entirely devoid of ACTH activity. Astwood, Raben and Payne (1952) have effected such a separation by adsorption of the active substance on cellulose or oxycellulose with subsequent elution, while in our laboratory we have similarly succeeded by passing a solution of the protein hormone ACTH through a column of the ion exchange resin Amberlite IRC-50 (Dixon, Moore, Stack-Dunne and Young, 1951). An active fragment almost certainly pre-exists in preparations of the protein in a form which is separable from the protein without hydrolysis, and the protein hormone probably has no significance, either biological or chemical. A protein which a few years ago appeared to be a homogeneous simple substance, responsible for an important hormonal action, has therefore been found not to be the hormone in question, and recent evidence suggests that the active material is a peptide of molecular weight about 3,000 (see Stack-Dunne and Young, 1954). How far then is assumption justified that growth hormone is itself diabetogenic or possesses others of the physiological actions given in Table 1? May not some or all of these activities properly be attributed to putative impurities lurking among the molecules of the protein itself? Raben and Westermeyer (1952) claimed that they had differentiated growth hormone from the pituitary factor which produces diabetes, but as the result of an exchange of samples between Drs. Raben and Westermeyer on the one hand, and Dr. Reid in my laboratory on the other, it was concluded by ourselves that Raben and Westermeyer's growth hormone was

indeed diabetogenic, though less so than ordinary growth hormone; but in our hands it was also less growth-promoting than other preparations of growth hormone. Some interesting biological variations associated with differences in the pH of the administered solution of growth hormone were found in these experiments, but such effects were in no way peculiar to the preparation of growth hormone by Raben and Westermeyer (see Young, 1953a).

TABLE 2. Procedures which Partially Inactivate but which Fail to alter significantly the Ratio of Diabetogenic Activity (cat) to Growth-promoting Activity (rat) for Ox Growth Hormone (Reid, 1952)

Repeated precipitation
Treatment with acid
Treatment with alkali
Acetylation
Iodination
Treatment with *B. subtilis* enzyme
Treatment with carboxypeptidase
Treatment with thioglycollic acid
Treatment with oxycellulose

In our laboratory we have taken the view that if the ratio of the diabetogenic activity to growth-promoting activity of preparations of growth hormone, subjected to various procedures which lead to partial inactivation, is always the same, then support is provided for the assumption that the same substance and the same active portion (or portions) of the same substance are responsible for both diabetogenic action and growth-promoting action. Such a finding does not prove identity, but it at least fails to reveal evidence against identity in circumstances where such evidence might well be revealed. In the experiments of this sort which Dr. Reid has carried out he has found that the ratio of diabetogenic activity in the cat to growth-promoting activity in the rat does not vary significantly from one preparation of growth hormone to another, nor is the ratio altered significantly by procedures designed to bring about partial inactivation of the hormone (Table 2). Evidence of a similar type was brought forward by Park *et al.* (1952), to the effect that the hypoglycaemic action of growth hormone in

hypophysectomized animals, as well as its anti-insulin effect, is due to growth hormone itself, but at present there is no similar evidence of which I am aware to support the view that the other actions described in Table 1 are indeed due to growth hormone *per se*. But equally there is no sound evidence against the view that some of these actions are properly ascribable to growth hormone.

Unfortunately, the chemical characterization of growth hormone as a protein is not entirely satisfactory. Differences are occasionally observed between the solubilities of different preparations of the hormone, and even the pH of minimum solubility is not always constant. The differences in solubility are not necessarily associated with differences in growth-promoting activity, and the reasons for them are uncertain. Reid (1951, 1952) has shown at Cambridge by the technique of Sanger that the N-terminal amino-acids in growth hormone are phenylalanine and alanine, and, in agreement with Li, that there are only two free α-amino groups per molecule of 47,000. Since there appear to be only two peptide chains in the molecule, a dissociation of growth hormone to smaller fragments of the order of 6,000 molecular weight, comparable with the dissociation of insulin, appears to be improbable. By the application of counter-current distribution to growth hormone Pierce (1954) found evidence for only one growth-promoting fraction in a number of preparations of growth hormone. Reid (1952) has further shown that partial acetylation such as to bring about complete acetylation of the free α-amino groups but incomplete acetylation of the ε-amino groups of lysine in the polypeptide chain, does not result in complete loss of growth-promoting action and diabetogenic activities, both these activities being diminished to the same extent. Furthermore, removal of the C-terminal amino-acids under the influence of carboxypeptidase does not result in loss of biological action (Reid, 1952). Although these fragments of evidence might be regarded as consonant with the view that the large molecule is not the biologically active material, the simplest assumption remains that growth hormone is indeed a protein, and that this protein is responsible for multiple biological actions. It appears desir-

able to sustain such a simple view until convincing evidence to the contrary is forthcoming.

The diabetogenic action of growth hormone is seen under somewhat restricted conditions. It is easily elicited in the normal intact adult dog and cat and less consistently in the rabbit. It is not seen in the lactating of pregnant adult female cat, nor in the intact rat. Arguing that diabetes is often a disease of old age we have administered very large doses of growth hormone to elderly senile male rats, but have found that they are able to respond to the growth-promoting action of the growth hormone without difficulty, and show no diabetes. Nevertheless, if the rat is partially depancreatized, the amount of pancreas removed being insufficient to produce diabetes of itself, the animal becomes susceptible to the diabetogenic action of growth hormone and such a partially depancreatized rat may respond to the administration of growth hormone by the exhibition of frank diabetes

It is perhaps of significance that carnivorous animals appear to be more susceptible to the diabetogenic action of growth hormone than non-carnivorous ones. This is comparable to the effects of surgical pancreatectomy on carbohydrate metabolism, where a much more severe diabetes is seen in carnivores than in herbivores. As Lukens has pointed out, the carnivorous animal is a hunting beast and commonly bolts its food. It is therefore much more dependent upon the availability of insulin for the rapid storage of the food which it has hurriedly absorbed than is the quietly and continuously feeding herbivorous animal. If growth hormone is regarded as a substance with an overall action in some respects antagonistic to that of insulin, it is not surprising that carnivorous animals should be more susceptible to the diabetogenic action of growth hormone than herbivorous ones.

[illegible]

College, London, have shown that growth hormone is capable of *antagonizing the hypoglycaemic action of insulin* in a patient with spontaneous hyperinsulinism (Black *et al.*, 1952), and there

is no reason to suppose that under suitable conditions treatment with growth hormone would not produce diabetes in human beings. In this respect, man stands somewhere between the carnivores and herbivores, in a position similar to that which he assumes with respect to the effects of surgical pancreatectomy on carbohydrate metabolism.

These differences in responsiveness to the diabetogenic action of growth hormone might lead us to expect that carnivorous animals would grow more effectively under treatment with growth hormone than would herbivores. In a limited series of experiments we have found differences of this sort. Certainly in some species of animal at any rate, growth is much more pronounced under the influence of growth hormone when the animal receives a high protein diet than when it receives a diet rich in carbohydrate. It is also of interest that the diabetogenic action of crude growth hormone is more effective in dogs receiving a high protein (meat) diet than in dogs receiving an isocaloric high carbohydrate diet (Young, 1949). Here we have an interesting parallelism between susceptibility to the growth-stimulating action of growth hormone and susceptibility to its diabetogenic action.

Growth Hormone and the Secretion of Insulin

Since growth hormone easily and rapidly induces diabetes on administration to an adult dog and cat it was of interest to investigate its action in a growing puppy or kitten. The experiments I am about to describe have all been carried out with rather crude preparations of growth hormone and, although it seems justified at the present time to ascribe the observed effects provisionally to growth hormone itself, a complete repetition of these experiments with the most purified growth hormone available is needed to be quite certain on this point.

When crude growth hormone was injected into a puppy or a kitten, the animal grew and did not become diabetic, even though the dose used would easily and rapidly produce diabetes in the adult animal (Young, 1941). In a proportion of the puppies which were treated for a long period with growth hormone diabetes did develop after six months or more, and as the

remain diabetic, though subsequent treatment with growth hormone induced diabetes without sustained growth. If then growth hormone and insulin were simultaneously administered to the animals, growth was resumed, but the amount of insulin required to combat the diabetogenic effect of the growth hormone was very large indeed. In other words, growth hormone had under these conditions induced an insulin-resistant diabetes (Young, 1944).

In a large number of experiments with kittens, although growth in response to the crude growth hormone was good, no diabetes ultimately developed in any one of the animals which had attained adulthood, although treatment was continued for a very long time. In this respect the dog and the cat appear to differ.

One may now ask whether, during the period of puppyhood and kittenhood, the pancreas of the animal was secreting enough insulin to combat the diabetogenic action of growth hormone, and whether the diabetogenic action of this hormone developed only when the pancreas was unable, for reasons at present undetermined, to furnish enough insulin to suppress the diabetes-inducing action of the hormone. I do not think it safe to assume that these growth-hormone treated young animals were indeed secreting such large amounts of insulin as are *found to be necessary to neutralize the diabetogenic effect of* the growth hormone when the animal had become adult. I should here like to recall that Langfeldt (1920) found that no diabetes developed in partially depancreatized puppies from which a sufficient proportion of the pancreas had been removed to induce severe diabetes in an adult dog. In one instance an enhanced sugar-tolerance was found as a result of removal of much of the pancreas from the puppy. After a time, however, as the animals grew a severe and ultimately fatal diabetes developed We unfortunately know very little about the changes in enzyme pattern which occur with normal growth, and it would be unwise to make any assumption about the constancy

of the sensitivity of the enzyme systems in the tissues to the action of growth hormone or other hormones during the period of growth. Nevertheless, the balance of evidence suggests that *the secretion of endogenous insulin appears to be an important* factor in the induction of growth under the influence of growth hormone. When the demands for the secretion of insulin exceed the capacity of the pancreas to respond, diabetes may develop in response to the administration of growth hormone.

I wish to refer to two recent pieces of evidence which agree with this view. In 1951, Milman, De Moor and Lukens (1951) showed that growth hormone was unable to induce nitrogen retention in the hypophysectomized depancreatized cat, and the same was true for the cat from which the pancreas alone had been removed. When growth hormone was given to a depancreatized cat maintained on a constant dosage of insulin some nitrogen retention was found under the influence of growth hormone, but maximum nitrogen retention, that is the nitrogen retention seen when the growth hormone was administered to a normal animal, was observed in the insulin-treated depancreatized cat only if the dosage of insulin administered was raised at the time the growth hormone was administered. Milman, De Moor and Lukens infer that the nitrogen-retaining action of growth hormone is normally dependent upon the *provision of an adequate supply of insulin by the pancreas.* More recently, Salter and Best (1953) have succeeded in inducing the hypophysectomized rat to grow substantially by treatment with insulin. Although the food intake was not controlled in these experiments a widening of the epiphyseal line was seen, and one must accept the conclusion of these authors that true growth had been induced in the hypophysectomized animal by insulin alone. We therefore have the paradoxical situation that growth hormone cannot induce growth in the absence of insulin, while insulin can nevertheless induce growth in the absence of growth hormone.

If growth hormone depends for its growth-promoting action on the secretion of insulin from the pancreas, then it is likely that the administration of growth hormone will be followed by a rise in the rate of secretion of insulin from the pancreas. *In*

an attempt to reveal such a rise, if it exists, by direct means, Bornstein, Reid and myself (1951) used alloxan-diabetic hypophysectomized adrenalectomized rats (ADHA rats) for the detection in insulin in the portal blood of growth-hormone treated cats. In these experiments we were unable to demonstrate the presence of a hypoglycaemic substance in the portal blood of the growth-hormone treated animal because such blood was found to exhibit a hyperglycaemic effect which, we inferred, was probably due to the presence of glucagon, secreted from the α-cells under the influence of growth hormone (see Table 1). More recently Dr. P. J. Randle and I, in unpublished investigations, have returned to the problem of the insulin content of the portal blood of the cat which has been treated with growth hormone, using the isolated diaphragm as the test-object, since the isolated diaphragm appears to be largely insensitive to any anti-insulin action of glucagon. We have so far failed to find unequivocal evidence for a rise in the insulin content of the portal blood as a result of growth hormone treatment, although these experiments are as yet far from complete and difficult to control. Dr. Randle (1954a, b) has also investigated the insulin activity, as assessed by the rat diaphragm method, of the blood of acromegalic patients and of patients with Simmonds' disease, finding that in acromegaly the insulin-like activity of the plasma is sometimes ten to fifteen times that of the normal person, while in some cases of Simmonds' disease it is about one-tenth of the normal value. If it is insulin alone which is being assayed by the isolated diaphragm in these experiments, Randle's results provide strong support for the view that under the influence of growth hormone an enhanced rate of liberation of insulin from the islets of Langerhans of the pancreas occurs But since growth hormone can itself exert an insulin-like action *in vitro* under some conditions (Table 1), Dr Randle has quite rightly refrained from assuming that a method of assay for insulin based on the *in vitro* response of the isolated rat diaphragm is specific for insulin when applied to the blood plasma of patients in whom the amounts of circulating *pituitary hormones may be abnormal.*

of the sensitivity of the enzyme systems in the tissues to the action of growth hormone or other hormones during the period of growth. Nevertheless, the balance of evidence suggests that the secretion of endogenous insulin appears to be an important factor in the induction of growth under the influence of growth hormone. When the demands for the secretion of insulin exceed the capacity of the pancreas to respond, diabetes may develop in response to the administration of growth hormone.

I wish to refer to two recent pieces of evidence which agree with this view. In 1951, Milman, De Moor and Lukens (1951) showed that growth hormone was unable to induce nitrogen retention in the hypophysectomized depancreatized cat, and the same was true for the cat from which the pancreas alone had been removed. When growth hormone was given to a depancreatized cat maintained on a constant dosage of insulin some nitrogen retention was found under the influence of growth hormone, but maximum nitrogen retention, that is the nitrogen retention seen when the growth hormone was administered to a normal animal, was observed in the insulin-treated depancreatized cat only if the dosage of insulin administered was raised at the time the growth hormone was administered. Milman, De Moor and Lukens infer that the nitrogen-retaining action of growth hormone is normally dependent upon the provision of an adequate supply of insulin by the pancreas. More recently, Salter and Best (1953) have succeeded in inducing the hypophysectomized rat to grow substantially by treatment with insulin. Although the food intake was not controlled in these experiments a widening of the epiphyseal line was seen, and one must accept the conclusion of these authors that true growth had been induced in the hypophysectomized animal by insulin alone. We therefore have the paradoxical situation that growth hormone cannot induce growth in the absence of insulin, while insulin can nevertheless induce [illegible] of growth hormone [illegible]

[illegible] action [illegible] s likely that the administration of growth hormone will be followed by a rise in the rate of secretion of insulin from the pancreas In

being with large babies and a diabetic diathesis. Miller, Hurwitz and Kuder (1944) were the first to associate large babies with a pre-diabetic condition and this relationship has been widely confirmed (cf. Gilbert and Dunlop, 1948). But there is at present no evidence for an association between the liability to produce large babies and an abnormally high secretion of growth hormone in these patients; furthermore the evidence that experimental treatment with growth hormone during pregnancy induces abnormally large offspring is not as yet conclusive (cf. Cotes, 1954).

When we turn to lactation the situation is a little clearer. The administration of growth hormone to cows in declining lactation leads to a significant increase in milk secretion (Cotes, Crichton, Folley and Young, 1949) and it is not unreasonable to associate this with the failure of growth hormone to induce diabetes in lactating bitches and cats. But again direct evidence that treatment with growth hormone increases the milk secretion of lactating bitches and cats is lacking as yet.

In pregnancy and lactation, as in the young growing animal, we may have circumstances in which the diabetogenic action of growth hormone is not demonstrable because the hormone is able, presumably with the co-operation of insulin secreted by the pancreas, to elicit a physiological response, namely foetal growth in pregnancy and extra milk secretion during lactation. The diabetes induced by growth hormone under some conditions can be regarded as the result of a failure of the mechanism by which the metabolites whose oxidation is depressed under the influence of the hormone are normally utilized for tissue [illegible]

of acromegaly, might lead to such a restraint upon the processes of carbohydrate oxidation, with an increase in appetite typical of growth hormone action, that diabetes ultimately develops Abaza and his colleagues (1953) have recently described a new condition under the name 'Syndrome de Young' in which prolonged growth, hyperlactation, diabetes, large babies, obesity and accidents of pregnancy are all associated. In such a condi-

The Mechanism of Action of Growth Hormone

In a tentative fashion we can view the intimate mechanism of [illegible] the [illegible] lets [illegible] rete glucagon. The latter substance assists to prevent the development of hypoglycaemia by stimulating hepatic glycogenolysis. Nevertheless the action of insulin can easily lead to dangerous hypoglycaemia, even in the presence of glucagon, particularly if the liver glycogen level is low, and growth hormone, or a substance derived from it (Bornstein and Park, 1953), depresses the sensitivity of the mechanism in the peripheral tissues through which insulin normally accelerates the uptake of glucose by the cell. This depression may be exerted directly on the enzyme hexokinase, or primarily on permeability barrier between hexokinase and the extracellular fluid and therefore only indirectly on the enzyme. There is then present in the blood stream a concentration of insulin considerably higher than normal but with a normal blood sugar level and little risk of hypoglycaemia; under these conditions the peripheral tissues, particularly muscle, will be stimulated to take up additional glucose and, either directly or indirectly, an additional supply of amino-acids. This uptake of amino-acids may be stimulated by growth hormone, or the substance derived from it, or by insulin, or, of course, by two or more factors acting together. As I have suggested elsewhere (Young, 1953b), the excess glucose utilized under these conditions may provide the additional energy needed for the deposition in the cell of unoxidized material, this deposition may then lead to cellular enlargement and possibly to mitosis.

Growth Hormone in Pregnancy and Lactation

I have referred above to the fact that growth hormone is not diabetogenic in pregnant or lactating bitches or cats (see Young, 1946). It might be thought that in the pregnant animal the administration of growth hormone should lead to the production of larger offspring. If indeed this is so, we might reasonably associate a slight oversecretion of growth hormone in the human

Reifenstein, E C. (Jr.), Kinsell, L. W. and Albright, F. (1946). *Endocrinology*, **39**, 71.

Salter, J. and Best, C. H. (1953). *Brit. med. J.* **2**, 353.

Stack-Dunne, M. P. and Young, F. G. (1954). *Ann. Rev Biochem.* **23** (in press).

[illegible]

[illegible]

p. 18.

tion we may have an exaggerated example of many entities which can all be associated with the action of growth hormone.

In growth hormone we may well have a substance whose physiological influence extends far beyond the field of action implied by its name.

REFERENCES

ABAZA, A., VARROUD-VIAL, J. and POMBAUTS, M. (1953). *Presse Médicale*, **61**, 495.

ASTWOOD, E. B., RABEN, M. S. and PAYNE, R. W. (1952). *Recent Prog. Horm. Res.* **7**, 1.

BLACK, K. O., MACDOUGALL, I., REID, E. and YOUNG, F. G. (1952), *Lancet*, **1**, 19.

[illegible]

[illegible] (1949). *Nature*, **164**, 992.

COTES, P. M., REID, E. and YOUNG, F. G. (1949). *Nature*, **164**, 209.

DIXON, H. B. F., MOORE, S., STACK-DUNNE, M. P. and YOUNG, F. G. (1951). [illegible]

[illegible]

LANGFELDT, E. (1920) *Acta Med Scand* **53**, 1.

MILLER, H. C., HURWITZ, D. and KUDER, K. (1944). *J Amer. med Ass.* **124**, 271.

MILMAN, A E., DE MOOR, P and LUKENS, F. D. W. (1951). *Amer. J. Physiol.* **166**, 354

PARK, C. R., BROWN, D. H., CORNBLATH, M, DAUGHADAY, W. H. and KRAHL, M E (1952). *J. biol. Chem* **197**, 151.

PIERCE, J. G (1954). *Biochem. J.* **57**, 16

RABEN, M. S. and WESTERMEYER, V. W. (1952) *Proc. Soc exp. Biol. Med.* **80**, 83

RANDLE, P. J. (1954a) *Lancet*, **441**, 809.

[illegible]

of disruption, of such family ties are productive of highly emotional reactions which in turn precipitate Graves' disease. The incidence of hyperthyroidism is said to decline in times of economic depression (Moersch, 1950) and it is possible that such social conditions bind families into closer units. In times of war, however, the incidence of Graves' disease may reach epidemic proportions in certain localities, and this is possibly to be correlated with the family disruption that occurs under such circumstances.

There is much evidence then that emotional stress may, in certain individuals, result in Graves' disease. The most likely chain of events underlying such a sequence may be represented: external stimulus → stimulation of central nervous system (emotional stress) → increased discharge of anterior pituitary thyrotrophic hormone (T.S.H.) → increased thyroid activity. Such a sequence of events has been studied experimentally by observing the effect of physical and emotional stress on thyroid activity in the rabbit (Brown-Grant, von Euler, Harris and Reichlin, 1953 a and b; Brown-Grant, von Euler, Harris and Reichlin, 1954, Brown-Grant, Harris and Reichlin, 1954), and the following account of the methods of measuring thyroid activity in this animal, and the effect of stress on such activity, is a summary of the results obtained by this group.

EXPERIMENTAL FINDINGS

(1) *Methods of measuring thyroid activity in the conscious rabbit*

Two methods are available for measuring thyroid activity in the conscious rabbit.

(*a*) *Rate of uptake of* ^{131}I *from the blood.* The radioactivity concentrated by the thyroid gland after intravenous injection of radioactive iodine has been measured by placing the rabbit, at various time intervals after injection, in an adjustable hammock with the ventral surface of its neck in a constant geometrical relationship with a Geiger-Muller tube In order to be able to perform repeated experiments in any one animal it is necessary to inject small doses of ^{131}I (about 2 μc.), and it then becomes necessary to approximate the Geiger-Müller tube closely to the animal's neck and to have a relatively wide side window in the

X

Stress and Thyroid Activity

G. W. HARRIS

Emotional Stress as an Aetiological Factor in Graves' Disease

Exophthalmic goitre was first described by Parry, a physician of Bath. He discussed, in lectures given in 1787 and later, a condition characterized by tachycardia, emotional excitement, exophthalmos and a swelling in the neck. Parry's description was not published until 1825, three years after his death. Further descriptions of this disease appeared later (Graves, 1835; Basedow, 1840). Both in these original accounts, and in many of the innumerable papers that have been published on Graves' disease since that time, reference is made to the fact that the patient often gives an account of an emotional shock shortly preceding the onset of the disorder. ... an aetio- ... ped con- ... r half of the nineteenth century, believed the disease was nervous in origin and that the swelling in the neck was a secondary manifestation.

More recently attempts have been made to analyse the type of personality prone to develop Graves' disease. Lidz and Whitehorn (1950), and others, have described such individuals as those highly dependent on interpersonal relationships for their security. This relationship is often familial, such as daughter to mother, or mother to child. The patient shows little adaptability and feels abject and helpless without such symbiotic relationship. It is thought that disruption, or threat

release curves observed in the present studies are outside the range of this possible error.

The slope of the release curve is proportional, apart from the slight error introduced by reaccumulation of ^{131}I, to the

$$\frac{\textit{amount of hormone secreted per unit time}}{\textit{amount of hormone in the gland}}$$

Providing therefore that the total amount of hormone in the gland remains constant, the slope of the curve may be taken as related to the amount of hormone secreted per unit time, that is to thyroid activity.

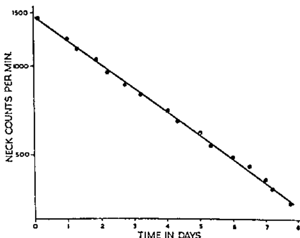


FIG 1 Normal release curve Rabbit injected with 2 μc. ^{131}I, and 48 hours later (that is at zero time on the abscissa) neck counts are started. Note that over the ensuing eight days the radioactivity of the thyroid decreases exponentially In this, and the following figures, the neck counts obtained have been corrected for physical decay of the ^{131}I (from Brown-Grant *et al*, 1954).

The advantages of using the release method of measuring thyroid activity is that it gives a measure of slow changes in thyroid activity, that it is less dependent on changes in the level of ^{127}I in the blood or on changes in renal excretion of iodine, and that each experiment is of sufficient duration to allow control and experimental periods of observation.

*lead shield of the tube. Under these circumstances the radio-*iodine in the blood vessels of the neck forms a proportion of the total neck counts. This proportion can be allowed for, and a 'thyroid count' can be extracted from the total neck counts. By observing the blood concentration of ^{131}I at various time intervals, values can then be obtained for the blood clearance of ^{131}I by the thyroid.

Brown-Grant and Gibson (1954) have recently improved this technique by using a scintillation counter and a rate-meter for obtaining an almost continuous record of thyroid radioactivity

release curves observed in the present studies are outside the range of this possible error.

The slope of the release curve is proportional, apart from the slight error introduced by reaccumulation of ^{131}I, to the

$$\frac{\textit{amount of hormone secreted per unit time}}{\textit{amount of hormone in the gland}}$$

Providing therefore that the total amount of hormone in the gland remains constant, the slope of the curve may be taken as related to the amount of hormone secreted per unit time, that is to thyroid activity.

[illegible] 1954).

The advantages of using the release method of measuring thyroid activity is that it gives a measure of slow changes in thyroid activity, that it is less dependent on changes in the level of ^{127}I in the blood or on changes in renal excretion of iodine, and that each experiment is of sufficient duration to allow control and experimental periods of observation.

(2) *The effect of various standard procedures on the slope of the release curve*

The effect of various procedures, well known to affect thyroid activity, was first tested on the slope of the release curve.

(*a*) *Hypophysectomy*. Removal of the pituitary gland results in a prompt and permanent decrease in the rate of release of ^{131}I. (Figure 2.) In a group of normal rabbits the average loss of radioactivity per day was 17·6 per cent, and after hypophysectomy 3·5 per cent.

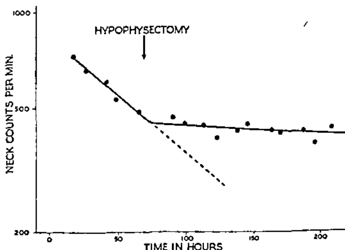
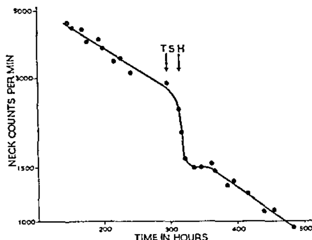


FIG. 2. To show the effect of hypophysectomy on the release curve. Repeated experiments on hypophysectomized rabbits show that the slow rate of release of ^{131}I is maintained indefinitely until the animal is killed (from Brown-Grant *et al.*, 1954).

(*b*) *Injection of thyrotrophic hormone*. Injection of thyrotrophic hormone (T.S.H.) into normal or hypophysectomized rabbits greatly increases the speed of release of radioactive hormone (Figure 3). The acceleration of thyroid activity starts within a few hours of administration of the hormone and persists for 10–20 hours. In the normal animal this period is followed by one in which the ^{131}I output of the thyroid is temporarily inhibited. This is probably due to the raised blood level of thyroid hormone 'feeding back' to inhibit T.S.H. secretion by the animal's own pituitary gland.



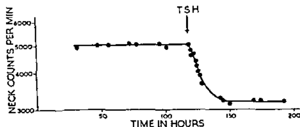


FIG. 3 *Above* To show the effect of injection of thyrotrophic hormone (T.S H) on the release curve of the normal rabbit *Below*. To show the effect of injection of thyrotrophic hormone (T.S.H.) on the release curve of the hypophysectomized rabbit (from Brown-Grant *et al* , 1954).

(*c*) *Injection of thyroxine.* Injection of thyroxine into normal rabbits promptly and markedly inhibits the release of ^{131}I from the thyroid. (Figure 4.) There are good grounds for believing that this action is due to the thyroxine inhibiting the release of pituitary T.S.H.

(*d*) *Exposure to cold.* In about half the experiments in which rabbits were taken from the constant temperature room (29°C.) and placed in a colder environment, a marked increase in the rate of output of ^{131}I from the thyroid gland was observed. (Figure 5.) It is of interest that the stimulus to the thyroid appeared greater if the animal were placed in an environment of 15–21°C., than if placed at 1°C. It is possible that the lower temperature may be acting as a non-specific stress, which inhibits thyroid activity (see below), as well as a specific stimulus.

From these results it seems clear that the various procedures which are known to result in definite effects on thyroid activity have the expected results on the thyroid release curve. It would seem valid then to interpret changes in the release curve, produced by other stimuli, in terms of thyroid activity.

(3) *The effect of emotional and physical stress on thyroid activity*

Many workers have reported that stresses of various types result in reduced ^{131}I uptake by the thyroid gland. Injection of typhoid vaccine, trauma, fasting, cold, heat (Williams, Jaffe and Kemp, 1949), injection of formalin (Paschkis, Cantarow, Eberhard and Boyle, 1950), cordotomy (Bogoroch and Timiras, 1951), anoxia, starvation (van Middlesworth and Berry, 1951), tourniquet shock (Hamolsky, Gierlach and Jensen, 1951), have all been reported to decrease the uptake of ^{131}I by the thyroid gland of the rat. However, measurements of ^{131}I uptake by the thyroid under conditions of stress do not necessarily give a specific index of thyroid activity, since stress (Bogoroch and Timiras, 1951) and the injection of cortisone (Ingbar, 1953) lead to an increased renal clearance of iodide, and may thus give an apparent decrease in thyroid uptake due to the fact that less ^{131}I is available to the thyroid.

The effect of emotional or physical stress has been studied

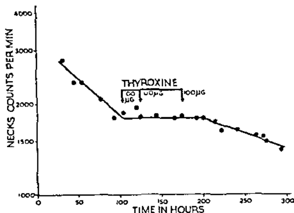


FIG. 4 To show the effect of thyroxine on the release curve of the normal rabbit. Thyroxine probably acts by inhibiting the release of thyrotrophic hormone from the anterior pituitary gland (from Brown-Grant *et al.*, 1954).

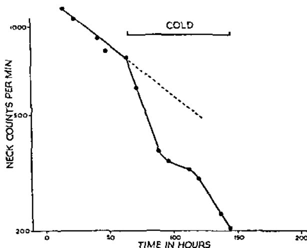


FIG. 5. To show the effect of a cold environment on the release curve of the normal rabbit (from Brown-Grant *et al.*, 1954).

on the rabbit. Normal rabbits have been subjected to a variety of stress stimuli during the course of a release curve. The emotional stimuli used have consisted of subcutaneous faradism, of subjecting the animals to restraint by tying their back legs to the side of their cage enabling them to sit but restricting their free movements, and of subjecting them to abrupt changes in illumination. The physical stresses have consisted of haemorrhage, laparotomy under ether anaesthesia, and intraperitoneal

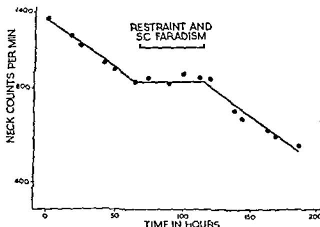


FIG. 6 To show the effect of emotional stress (induced by restraint and subcutaneous faradism) on the release curve of the normal rabbit (from Brown-Grant *et al*, 1954).

injections of turpentine In the great majority of experiments these stress stimuli resulted in a prompt and marked (in many cases complete) inhibition of the release of thyroidal radioactivity (Figure 6). Animals exposed to a constant emotional stress often showed a return of the slope of the release curve to that of the control period before the cessation of the stress. This 'escape' is most likely due to the animals becoming 'accustomed' to the stressing procedure, since the overt behaviour of the animals changes, suggesting that the procedure had become less disturbing. Furthermore it was found that alternation of two emotional stress stimuli would prolong the inhibitory effect on the thyroid, as compared with a single constant stimulus.

The interpretation that a decreased slope of the release curve

of the stressed rabbit is related to decreased release of thyroid hormone is supported by the finding that the blood level of organically-bound ^{131}I is simultaneously reduced (Brown-Grant, 1954).

The observation that emotional stress results in decreased thyroid activity in the rabbit is contrary to the results of Kracht and Kracht (1952). These workers reported that wild rabbits brought into captivity develop a state of hyperthyroidism (exophthalmos, weight loss and muscular weakness) and die within three weeks. Their criteria of a hyperactive thyroid, however, were indirect and histological, rather than actual measurement of thyroid activity. Repetition of these experiments (Brown-Grant, Harris and Reichlin, 1954), on a small group of wild rabbits, failed to confirm these findings. Wild rabbits were noted to live and maintain body weight during laboratory existence, and their output of thyroidal ^{131}I was seen to react in the same way as that of the domestic animal, on subjection to emotional stresses.

(4) *The mechanism of the thyroid inhibition to stress*

Several possible mechanisms underlying the inhibitory response of the thyroid T
may be, (*a*) mediated b
gland, (*b*) secondary to
cortex, (*c*) determined by a decreased release of pituitary thyrotrophic hormone, or (*d*) dependent on some other mechanism. Experiments have been conducted to test these possibilities.

(*a*) *Nerve supply to the thyroid gland.* There is no good evidence that the thyroid gland is supplied by secretomotor nerve fibres (see Harris, 1948). Bilateral removal of the stellate ganglia and a portion of the cervical sympathetic chain had no apparent effect on the thyroid response to stress (Figure 7).

(*b*) *The adrenal gland.* Injection of adrenaline (in large doses, 500–1,000 μg./day for two to six days), cortisone and adrenocorticotrophic hormone was found to inhibit the release of thyroidal ^{131}I in the rabbit. In view of this, and the fact that activation of the adrenal medulla and cortex is well known to follow stress stimuli, it seemed possible that the thyroid response

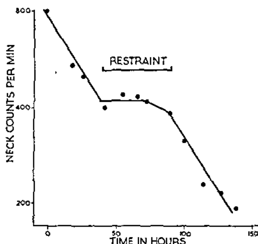


FIG. 7. To show the effect of emotional stress (induced by restraint) on the release curve of a rabbit previously subjected to removal of the stellate ganglia and a portion of the cervical sympathetic chain (from Brown-Grant *et al.*, 1954).

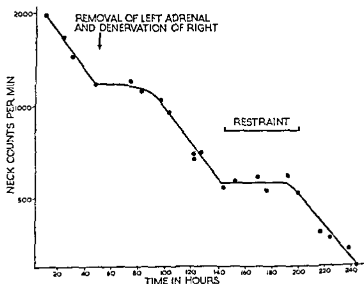


FIG. 8 To show the effect of emotional stress (restraint) on the release

was secondary to increased liberation of adrenaline and/or adrenal cortical hormone. However, animals with the adrenal glands denervated (Figure 8), or removed (maintained on constant cortisone regime) (Figure 9) showed, in the majority of experiments, thyroid inhibitory responses to stress stimuli as in normal rabbits. It may be argued that the rate of utilization of cortisone under conditions of stress is unknown, but it is unlikely that it is decreased.

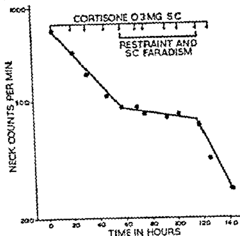


FIG. 9 To show the effect of emotional stress (restraint and subcutaneous faradism) on the release curve of an adrenalectomized rabbit, maintained on constant cortisone therapy (0 3 mg. s c , twice daily) (from Brown-Grant *et al.*, 1954).

(*c*) *Pituitary thyrotrophic secretion.* It is difficult to investigate the effect of stress on thyroid activity in the hypophysectomized rabbit. In most hypophysectomized animals the rate of release of thyroid radioactivity is so low that it would be difficult to detect a significant slowing under the effect of a stress stimulus.

One recent observation makes the pituitary origin of the stress inhibition of the thyroid probable. A series of rabbits in which the pituitary stalk had been divided and a waxed paper plate inserted between the cut ends of the stalk, have been tested for the stress response (Brown-Grant, Harris and Reichlin, unpublished) These experiments are still in progress, but it appears probable that stalk section abolishes the thyroid res-

ponse to emotional stress but not to physical trauma. The interpretation provisionally placed on these results is that emotional stress affects thyroid activity by an action through the hypothalamus and hypophysial portal vessels, inhibiting the pituitary release of T.S.H. Physical trauma, on the other hand, seems to inhibit T.S.H. secretion by a direct action on the pituitary gland, possibly via tissue breakdown products carried by the blood stream.

Relationship of Present Findings to Aetiology of Graves' Disease

The present findings indicate that thyroid function is affected by the central nervous system, and that emotional stress may cause *rapid and marked changes in thyroid activity*. In view of the many accounts in the clinical literature that emotional stress often precedes the pathological increase in activity of the thyroid, known as Graves' disease, it is perhaps surprising to find that emotional stimuli result in thyroid inhibition in the normal rabbit.

The responses of the adrenal cortex and the thyroid gland of the rabbit to stress appear to bear a reciprocal relationship, the adrenal cortex being activated whilst the thyroid is inhibited. This finding recalls to mind the older views of Marine, and others, that the hyperactive thyroid of the Graves' patient is to be correlated with atrophy of the adrenal cortex. In this respect the following observations appear to be relevant.

respiratory exchange, and a symptom complex resembling exophthalmic goitre (Marine and Baumann, 1921).

(*ii*) Thyrotoxicosis is often associated with signs of adrenocortical under-activity (large thymus, lymphoid hyperplasia) and a small adrenal cortex (Marine, 1930).

(*iii*) Graves' disease has been noted to follow X-ray damage to the adrenal cortex (Oppenheimer, 1937), and the incidence of this condition has been reported as ten times greater in patients suffering from Addison's disease than in normals (Frederickson, 1951).

(*iv*) Patients suffering from Cushing's syndrome may show signs of hypothyroidism (high level of serum cholesterol, dry skin, and low basal metabolic rate) (Heinbecker, 1944), and an inactive thyroid (Cushing, 1933).

It would seem that there are good reasons for believing that the adrenal cortex is in some way involved in Graves' disease. Whether the atrophy of the adrenal cortex is involved from the aetiological point of view, or whether it is a consequence and part of the developed condition, remains to be seen.

An interesting speculation is that the patient with Graves' disease has responded to stress with the opposite response of the normal. That is, that some stress stimulus has resulted in increased thyroid, and decreased adrenal cortical, activity. Such a view would suggest that some neurological mechanism is primarily at fault, and would indicate that the condition is really a 'disease' rather than the result of an exaggerated physiological response.

REFERENCES

VON BASEDOW, K. A (1840) *Wsch. f. d ges Heilk.* (Casper), **197**, 220.
BOGOROCH, R. and TIMIRAS, P (1951). *Endocrinology*, **49**, 548
BROWN-GRANT, K (1954) Unpublished.
BROWN-GRANT, K, VON EULER, C, HARRIS, G. W. and REICHLIN, S. (1953a). *Memoirs of the Society for Endocrinology*, **1**, 29.
BROWN-GRANT, K, VON EULER, C, HARRIS, G W. and REICHLIN, S. (1953b) *J Endocrinol* **9**, xliii.
BROWN-GRANT, K, VON EULER, C, HARRIS, G. W. and REICHLIN, S (1954) *J Physiol.* **126**, 1
BROWN-GRANT, K, HARRIS, G W and REICHLIN, S (1954) *J. Physiol* **126**, 29.
BROWN-GRANT, K. and GIBSON, J. G (1954). *J Physiol.* in press.
CUSHING, H (1933) *Arch Int Med* **51**, 487.
FREDERICKSON, D S (1951). *J clin Endocrinol* **11**, 760.
GRAVES, R J (1835) *London Medical & Surgical Journal*, part 2, published by Renshaw, **7**, 516
HAMOLSKY, M W, GIERLACH, Z S. and JENSEN, H (1951). *Amer J. Physiol.* **164**, 35.
HARRIS, G W. (1948) *Physiol Rev* **28**, 139.

HEINBECKER, P. (1944). *Medicine*, **23**, 225.

INGBAR, S. H. (1953). *Endocrinology*, **53**, 171.

KRACHT, J. and KRACHT, URSULA (1952). *Virch. Arch.* **321**, 238.

LIDZ, T. and WHITEHORN, J. C. (1950). *Res. Publ. Ass. nerv. ment. Dis* **29**, 445.

MARINE, D. (1930). *Am. J. Med. Sci.* **180**, 767.

MARINE, D. and BAUMANN, E. J. (1921). *Amer. J. Physiol.* **57**, 135.

VAN MIDDLESWORTH, L. and BERRY, M. M. (1951). *Amer. J. Physiol.* **167**, 576.

[illegible]

PASCHKIS, K. E., CANTAROW, A. H., EBERHARD, T. and BOYLE, D. (1950). *Proc. Soc. exp. Biol., N.Y.* **73**, 117.

WILLIAMS, R. H., JAFFE, H. and KEMP, CAROL (1949). *Amer. J. Physiol* **159**, 291.

XI

The Physiological Actions of the Sex Hormones

P. M. F. BISHOP

SIR HENRY DALE, in the inaugural lecture[1] of this series, drew attention to two striking developments, first the [illegible] of medicine and science has accumulated during the present century. The subject of this lecture is no exception to this trend. For practical purposes I can truthfully say that all the important discoveries concerning the action of the sex hormones have not only been made during my lifetime, but during the years in which I have been actively engaged in the study of endocrinology. Furthermore, in the time at my disposal, I am at a loss to know, not what to say, but what to leave out, for I cannot hope to discuss all the physiological actions of the sex hormones. I therefore propose to confine myself to three aspects only of this very extensive problem; and accordingly intend to introduce the subject with a short historical review in which I hope to indicate the contribution to our present knowledge that has been supplied by biologists and biochemists. Next I will deal in some detail with the problem of pituitary-gonadal relations. Finally, I shall give a brief account of the physiological action of androgens.

[1] Not included in the present volume.

HISTORICAL

From the third century B.C. until the first half of the seventeenth century A.D. the ovary was referred to as the female testis. Regnier de Graaf was one of the first to give a detailed and accurate account of the ovaries, in 1672, and is said to have died at the early age of thirty-two of indignation resulting from a dispute concerning his claims to priority for his observations on the generative organs. Haller (1766) spoke of the ovarian follicles as the 'Graafian vesicles'. Fallopius (1561) first observed the corpus luteum, though de Graaf gave the first detailed account of it in 1672. Malpighi in 1686 named it the 'corpus luteum'. Fifty years ago Fraenkel (1903) established the function of the corpus luteum in relation to reproduction.

In 1651 Nathaniel Highmore gave an account of the testis and in 1668 de Graaf made a more detailed study of its structure. In 1849 Berthold grafted testes into capons and restored their cock's combs, and in 1889 Brown-Séquard rejuvenated himself to his own satisfaction by means of injections of a mixture of the juice of testes, the blood of the spermatic vein and water.

The modern era in the study of the sex hormones was opened in 1923 when Allen and Doisy reintroduced the vaginal smear technique, originally described by Stockard and Papanicolaou in 1917, as a convenient and rapid method of biological assay of oestrogenic material. In 1927 Aschheim and Zondek, besides laying the foundation of the biological tests for pregnancy, demonstrated the presence of large amounts of an oestrogenic principle in the urine of pregnant women, and in the same year McGee showed that there was an androgenic substance in bull's testis. In the following year Loewe and his colleagues extracted an androgenic substance from male urine. In the same year, 1928, Corner and Willard Allen demonstrated that the corpus luteum contains a principle which can produce progestational changes in the uterus. In 1929, Marrian isolated what was later to become known as pregnanediol from the urine of pregnant women. In the latter months of this year and in the early months of 1930, oestrone, as it was later called, was isolated almost simultaneously by Doisy and his group (1929) in St. Louis, Butenandt (1929) in Germany, and Dingemanse and her

colleagues (1930) in Holland, and another oestrogen, later known as oestriol, was isolated by Marrian (1930). In 1931 Kober elaborated a colorimetric method of identifying oestrogenic compounds. Also in 1931 Butenandt and Tscherning extracted 15 mg. of a pure crystalline substance which they called 'androsterone' from 15,000 litres of male urine.

All these achievements were especially remarkable because at this time there was no clear conception of the structural formulae of these compounds. The brilliant studies of Windaus and his school (1928) and of Wieland and his collaborators (1928) had gone far to associate the sterols, the bile acids and the cardiac glucosides into a single group, but the structural formulae of cholesterol and ergosterol were still tentative. In 1932 Rosenheim and King re-examined all the available evidence and identified the basic structure of these compounds with the now familiar cyclopenteno-phenanthrene ring system. This was the second milestone in the modern era of the history of the sex hormones.

Following this event, discovery proceeded apace. Commencing in the same year, Girard and his colleagues (1932, 1936), in Paris, extracted from the urine of pregnant mares large quantities of oestrone, equilin and equilenin, and Schwenck and Hildebrandt (1932) partially synthesized oestradiol by hydrogenation of oestrone. In 1934 Cook and his colleagues (Cook and Girard, 1934; Cohen *et al.*, 1934, 1935) completely elucidated the structure of the natural oestrogens, and in the same year Ruzicka and his group in Switzerland identified Butenandt's 'androsterone' as aetio*allo*cholan-3 (α)-ol-17-one. Still in the same year Butenandt and Dannenbaum isolated another androgen from male urine, which eventually proved to be dehydroisoandrosterone. Finally, in this same year, 1934, the isolation of pure crystalline progesterone was announced by four groups of workers, Butenandt, Willard Allen and Wintersteiner, Slotta and his colleagues, and Hartmann and Wettstein. Incidentally it is interesting to note that Slotta was Fraenkel's son-in-law. In 1935 Doisy (MacCorquodale *et al.*, 1935) succeeded in extracting a few milligrammes of pure crystalline oestradiol (the true follicular hormone), obtained

Historical

From the third century B.C. until the first half of the seventeenth century A.D. the ovary was referred to as the female testis. Regnier de Graaf was one of the first to give a detailed and accurate account of the ovaries, in 1672, and is said to have died at the early age of thirty-two of indignation resulting from a dispute concerning his claims to priority for his observations on the generative organs. Haller (1766) spoke of the ovarian follicles as the 'Graafian vesicles'. Fallopius (1561) first observed the corpus luteum, though de Graaf gave the first detailed account of it in 1672. Malpighi in 1686 named it the 'corpus luteum'. Fifty years ago Fraenkel (1903) established the function of the corpus luteum in relation to reproduction.

In 1651 Nathaniel Highmore gave an account of the testis and in 1668 de Graaf made a more detailed study of its structure. In 1849 Berthold grafted testes into capons and restored their cock's combs, and in 1889 Brown-Séquard rejuvenated himself to his own satisfaction by means of injections of a mixture of the juice of testes, the blood of the spermatic vein and water.

The modern era in the study of the sex hormones was opened in 1923 when Allen and Doisy reintroduced the vaginal smear technique, originally described by Stockard and Papanicolaou in 1917, as a convenient and rapid method of biological assay of oestrogenic material. In 1927 Aschheim and Zondek, besides laying the foundation of the biological tests for pregnancy, demonstrated the presence of large amounts of an oestrogenic principle in the urine of pregnant women, and in the same year McGee showed that there was an androgenic substance in bull's testis. In the following year Loewe and his colleagues extracted an androgenic substance from male urine. In the same year, 1928, Corner and Willard Allen demonstrated that the corpus luteum contains a principle which can produce progestational changes in the uterus. In 1929, Marrian isolated what was later to become known as pregnanediol from the urine of pregnant women. In the latter months of this year and in the early months of 1930, oestrone, as it was later called, was isolated almost simultaneously by Doisy and his group (1929) in St. Louis, Butenandt (1929) in Germany, and Dingemanse and her

Pituitary-Gonadal Relations

The identity of the various elements of the pituitary gonadotrophic complex remained in doubt for some time, but now it is certain that there are at least two hormones which have been chemically isolated and purified, though not yet structurally characterized. They have unfortunate names which are historical relics rather than accurately descriptive designations in the light of modern knowledge of their functions. In the late 'twenties and early 'thirties when the work of Aschheim and Zondek (Aschheim and Zondek, 1928; Zondek, 1930) stimulated studies of the response of the ovary of the immature mouse and rat to crude gonadotrophic extracts, it became clear that there were two distinct reactions—the production of multiple enlarged follicles and the development of relatively few mature corpora lutea. The hypothetical hormones responsible for these reactions were therefore referred to, respectively, as the follicle-stimulating hormone (FSH) and the luteinizing hormone (LH). But, of course, these hormones have equally important functions in relation to the testis. The so-called luteinizing hormone stimulates the interstitial cells of Leydig to produce testosterone, and the Americans, quite logically, refer to it as the interstitial-cell-stimulating hormone (ICSH) (Evans and Simpson, 1950), for this would apply also to the interstitial theca-lutein cells which are transformed into the corpus luteum, and produce progesterone. No new name, however, has been coined for FSH, which has no follicles to stimulate in the male gonad, but is responsible for ensuring the orderly reduction-division of the germinal epithelium in the process of spermatogenesis.

Figures 1–6 indicate simply and briefly the main events in the pituitary-ovarian cycle, and I pass on to consider in more detail the evidence for these pituitary-gonadal relations (Greep and Jones, 1950a), obtained by the use of such methods as the following.

(*i*) Injection of gonadal hormones into pubescent male and female rats.

(*ii*) Injection of gonadal hormones into castrated or spayed pubescent rats.

(*iii*) Study of body and organ weights.

from sow's ovaries. In the same year Laqueur (David *et al.*, 1935), in Amsterdam, isolated pure crystalline testosterone from the testis. In 1935 and 1936 Zimmermann elaborated a colorimetric, quantitative method of determining androgenic substances possessing a ketone group at the 17-carbon atom. In 1936 Girard announced a series of reagents by which it is possible to separate the ketonic from the non-ketonic fractions of steroid substances from urine. In 1936, Venning and Browne, in Montreal, elaborated a gravimetric method of estimating pregnanediol glucuronide from the urine of pregnant women, as a metabolite of progesterone.

By 1938 E. C. Dodds (now Sir Charles) and his colleagues at the Middlesex Hospital had completed a laborious but fascinating search for a potent synthetic oestrogen. In 1933, Cook, Dodds and Hewett observed that the compound 1-keto-1,2,3,4-tetrahydrophenanthrene possessed oestrogenic activity. This led to a study of other hydrocarbons, and it appeared that the phenanthrene structure was not essential for oestrogenic activity. Dodds then set out to find a simpler, and yet potent, compound. Eventually the very simple compound anol(p-hydroxy-propenyl-benzene) appeared to be very highly active. Repetition of the work, however, did not confirm the original results, and it seemed clear that the potent oestrogenic activity was due to a contaminant. Further studies, in collaboration with Sir Robert Robinson, revealed that this potent oestrogen was in fact a condensation product of two molecules of anol, namely 4:4'-dihydroxy-α:β-diethylstilbene, which Dodds christened 'stilboestrol'. Here was an oestrogen, easy to make in the laboratory, and highly potent even when given by mouth. It was the first example of a synthetic compound being more potent than its naturally occurring equivalent. In the same year, 1938, Dr. and Mrs. Callow and Dr. Emmens developed a reliable and reasonably simple method for chemical estimation of the urinary neutral 17-ketosteroids. Two years later Bachmann and his colleagues at Ann Arbor completed the total synthesis of equilenin, and in 1948, Anner and Miescher achieved the total synthesis of oestrone.

Pituitary-Gonadal Relations

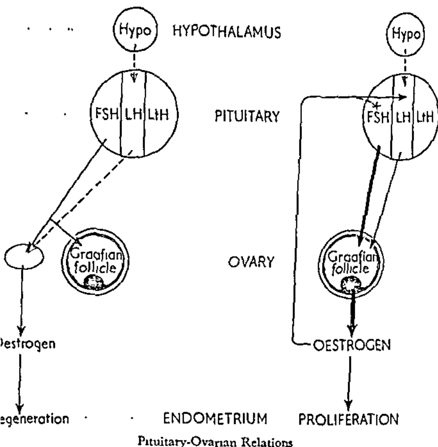
The identity of the various elements of the pituitary gonadotrophic complex remained in doubt for some time, but now it is certain that there are at least two hormones which have been chemically isolated and purified, though not yet structurally characterized. They have unfortunate names which are historical relics rather than accurately descriptive designations in the light of modern knowledge of their functions. In the late 'twenties and early 'thirties when the work of Aschheim and Zondek (Aschheim and Zondek, 1928; Zondek, 1930) stimulated studies of the response of the ovary of the immature mouse and rat to crude gonadotrophic extracts, it became clear that there were two distinct reactions—the production of multiple enlarged follicles and the development of relatively few mature corpora lutea. The hypothetical hormones responsible for these reactions were therefore referred to, respectively, as the follicle-stimulating hormone (FSH) and the luteinizing hormone (LH). But, of course, these hormones have equally important functions in relation to the testis. The so-called luteinizing hormone stimulates the interstitial cells of Leydig to produce testosterone, and the Americans, quite logically, refer to it as the interstitial-cell-stimulating hormone (ICSH) (Evans and Simpson, 1950), for this would apply also to the interstitial theca-lutein cells which are transformed into the corpus luteum, and produce progesterone. No new name, however, has been coined for FSH, which has no follicles to stimulate in the male gonad, but is responsible for ensuring the orderly reduction-division of the germinal epithelium in the process of spermatogenesis.

Figures 1–6 indicate simply and briefly the main events in the pituitary-ovarian cycle, and I pass on to consider in more detail the evidence for these pituitary-gonadal relations (Greep and Jones, 1950a), obtained by the use of such methods as the following:

(*i*) Injection of gonadal hormones into pubescent male and female rats.

(*ii*) Injection of gonadal hormones into castrated or spayed pubescent rats.

(*iii*) Study of body and organ weights.



Pituitary-Ovarian Relations

FIG. 1. During the menstrual phase.

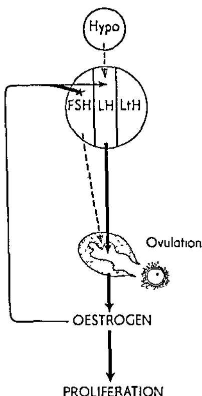
FIG. 2. In the proliferative phase.

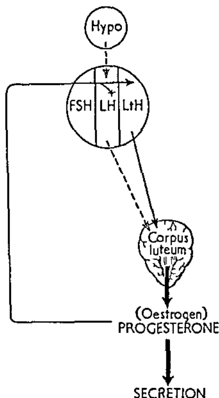
(*iv*) Histological and histochemical studies of section of the pituitary, gonads, uterus, vagina and adrenals, with special reference to cholesterol content.

(*v*) Noting the time of opening of the vaginal orifice and study of vaginal smears.

(*vi*) Injections of pituitaries of rats submitted to experimental conditions into immature hypophysectomized female rats, to demonstrate the gonadotrophin content of the pituitaries, by studying the effects on the ovaries and uteri of the recipient rats.

(*vii*) Autotransplantation of ovaries into the spleen, so that





Pituitary-Ovarian Relations

FIG 3 At ovulation.

FIG. 4. At the height of corpus luteum development.

the ovarian hormones are inactivated in the liver. The ovaries may, however, be stimulated by pituitary gonadotrophins.

(*viii*) Parabiotic studies with one member of the union gonadectomized and the other hypophysectomized.

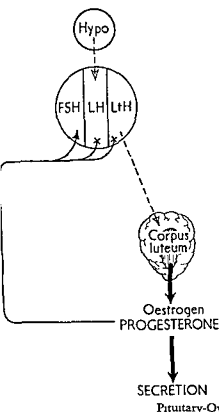
(*ix*) Direct stimulation, or section of, or application of drugs to, the hypothalamus, neuro-hypophysial pathways, or the pituitary gland.

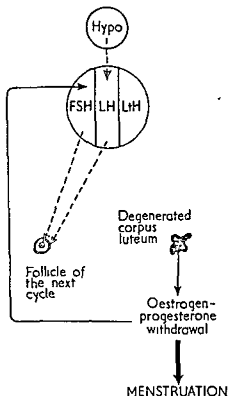
In the first place there is a fundamental difference between the pituitary-testicular relations and the pituitary-ovarian relations. In the male LH stimulates the interstitial cells of Leydig

to produce testosterone, and a balance is struck between the production of the pituitary and testicular hormone which maintains a constant output of androgen. FSH stimulates the seminiferous tubule and, in normal circumstances, spermatogenesis proceeds at a remarkably constant rate. Whether there is a second testicular hormone which moderates the output of FSH is a matter of controversy. It is a fact that in castrates and eunuchoids and in cases of seminiferous tubule failure the urinary output of FSH is considerably elevated, and it has been postulated therefore that there is a pituitary moderator, which some refer to as 'inhibin' (McCullagh, 1932), elaborated probably by the sustentacular cells of Sertoli (Howard *et al.*, 1950). Others, on the other hand, believe that the germinal epithelium 'utilizes' or consumes the FSH which stimulates it, and therefore when the germinal epithelium is lacking, 'non-utilization' of FSH leads to its circulation and excretion in raised concentration (Heller *et al.*, 1953). This utilization hypothesis would be in harmony with theories concerning the thyroid-stimulating hormone (TSH) and the thyroid cells. It is supposed that TSH sets in motion the oxidase systems which control the rate at which the thyroid cell manufactures its hormone, and that in doing so TSH is itself inactivated. Luteotrophic hormone has been identified in the male pituitary gland (of the rat) but so far no function has been associated to it (Greep and Jones, 1950b).

In contrast to these relatively simple arrangements in the male the pituitary-ovarian relations consist of a series of sequential percussions. FSH is responsible for growth of the follicle, of the granulosa cells and of the theca interna, but not for the secretion of oestrogen (Greep *et al.*, 1942). LH primes the corpus luteum, probably by helping it to store cholesterol (Everett, 1947), but is not responsible for the secretion of oestrogen or progesterone (Greep, 1938). As Hisaw (1947) puts it, the ovarian tissue must be made 'competent' before its hormones can be secreted.

FSH plus LH, however, leads to hypertrophy of the theca interna (which is currently supposed to be the normal source of the follicular hormone, oestradiol) and secretion of oestrogen (Greep *et al.*, 1942).





Pituitary-Ovarian Relations

FIG 5 At the beginning of corpus luteum development.

FIG. 6 At the commencement of menstruation.

The short-term effect, possibly in terms of low dosage, of oestrogen on the pituitary is to stimulate a burst of FSH output followed immediately by diminished secretion and storage of hormone. If the production of FSH gains dominance then the pituitary has little ability to produce LH. But the burst of FSH leads to more oestrogen being secreted, and thus to progressive suppression of FSH production with consequent release of LH. Thus the reversal of the FSH/LH ratio in favour of LH induces ovulation. There is some evidence that progesterone can assist oestrogen in releasing LH (Everett, 1948). Hooker and Forbes

(1947) elaborated a technically superb method of micro-assay of progesterone involving recording the changes in the stromal nuclei in a minute but accurately measured strip of the spayed mouse's endometrium, following injection of the extract to be assayed into this strip. This bio-assay method can identify one five-millionth of a milligramme of progesterone. By means of it they have identified progesterone in appreciable quantities in the blood of women just before ovulation. Progesterone is of course secreted under the influence of Lt.H, though there is no indication as to how Lt.H appears on the scene in this pre-ovulatory phase, but when it assumes dominance in the pituitary gonadotrophic complex, there is little ability to produce FSH and LH. Though progesterone may enable oestrogen to produce LH before ovulation, after this event has taken place, progesterone inhibits LH production. Thus the stage is set for Lt.H to maintain the corpus luteum and its secretion of progesterone. *Perhaps it is the failure, in the late post-ovulatory phase,* to elaborate FSH and LH, which leads to oestrogen and to progesterone withdrawal and the onset of menstrual bleeding in women and primates.

Lt.H has been identified with prolactin in rats, but so far not in the human species (Astwood, 1941). Injection of prolactin in women has not prolonged the luteal phase of the cycle (Bradbury *et al.*, 1950). On the other hand it has been established that if human chorionic gonadotrophin, with properties resembling LH rather than Lt.H, is injected in daily doses of 5,000 to 10,000 I.U. not later than the twenty-first day of the cycle, it will prolong the cycle by maintaining the corpus luteum (Bradbury *et al.*, 1950).

Oestrogen leads to storage of cholesterol in all ovarian tissues and especially in the corpus luteum. It is suggested by one group of workers that LH also causes cholesterol to be stored, though others postulate that LH releases the stored cholesterol from the corpus luteum and that Lt.H converts it into progesterone (Everett, 1947). Whereas oestrogen is a cholesterol 'enhancer', androgen is a cholesterol 'depletor' (Greep and Jones, 1950a).

The effects of oestrogen and androgen on the pituitary are very similar. They both increase the number of eosinophil cells,

and decrease the number of basophil cells, at the same time causing degranulation of these latter cells. Oestrogen, in addition, leads to enlargement of the gland owing to hyperplasia of the chromophobe cells (Severinghaus, 1937; van Dyke, 1936,

[illegible]

LH production, continued or high dosage administration suppresses all gonadotrophic activity by exhausting the stores and preventing the cells from elaborating fresh hormone (Leonard *et al.*, 1931; Meyer *et al.*, 1932).

In the absence of oestrogen production by the ovary FSH is both stored and secreted in abundance, and thus high concentrations of the hormone are found in the urine of menopausal women and eunuchoid men. Administration of oestrogen in such cases immediately suppresses FSH storage and production, and the urinary FSH values diminish. Administration of androgen, however, has little effect on FSH production (McCullagh *et al.*, 1948), for it has already been pointed out that androgen is balanced against LH and not against FSH.

The hypothalamus and neurohumoral stimuli to the pituitary play an important role in the pituitary-gonadal relations. Neurogenic impulses are carried from the paraventricular and supra-optic nuclei along the nerve fibres to the region of the median eminence, whence the hypophysial portal circulation conveys humoral impulses to the zona tuberalis of the anterior lobe (Harris, 1948). The cell population of this zone undergoes kaleidoscopic changes with corresponding alterations in gonadotrophin secretion as the result of stimulation or lesions of these hypothalamic centres (Dawson, 1948). The release of LH and the induction of ovulation have been studied in this connection. The main impulse seems to be adrenergic, for adrenaline instilled into the pars distalis of the rabbit's pituitary leads to ovulation (Markee *et al*, 1948), and dibenamine—an anti-adrenergic drug—if injected within one minute of copulation will block the normal coitus-induced ovulation (Sawyer *et al.*, 1947). There is probably a preceding cholinergic process, for atropine administered at a similar time in relation to coitus often

(1947) elaborated a technically superb method of micro-assay of progesterone involving recording the changes in the stromal nuclei in a minute but accurately measured strip of the spayed mouse's endometrium, following injection of the extract to be assayed into this strip. This bio-assay method can identify one five-millionth of a milligramme of progesterone. By means of it they have identified progesterone in appreciable quantities in the blood of women just before ovulation. Progesterone is of course secreted under the influence of Lt.H, though there is no indication as to how Lt.H appears on the scene in this pre-ovulatory phase, but when it assumes dominance in the pituitary gonadotrophic complex, there is little ability to produce FSH and LH. Though progesterone may enable oestrogen to produce LH before ovulation, after this event has taken place, progesterone inhibits LH production. Thus the stage is set for Lt.H to maintain the corpus luteum and its secretion of progesterone. Perhaps it is the failure, in the late post-ovulatory phase, to elaborate FSH and LH, which leads to oestrogen and to progesterone withdrawal and the onset of menstrual bleeding in women and primates.

Lt.H has been identified with prolactin in rats, but so far not in the human species (Astwood, 1941). Injection of prolactin in women has not prolonged the luteal phase of the cycle (Bradbury *et al.*, 1950). On the other hand it has been established that if human chorionic gonadotrophin, with properties resembling LH rather than Lt.H, is injected in daily doses of 5,000 to 10,000 I.U. not later than the twenty-first day of the cycle, it will prolong the cycle by maintaining the corpus luteum (Bradbury *et al.*, 1950).

Oestrogen leads to storage of cholesterol in all ovarian tissues and especially in the corpus luteum. It is suggested by one group of workers that LH also causes cholesterol to be stored, though others postulate that LH releases the stored cholesterol from the corpus luteum and that Lt.H converts it into progesterone (Everett, 1947). Whereas oestrogen is a cholesterol 'enhancer', androgen is a cholesterol 'depletor' (Greep and Jones, 1950a).

The effects of oestrogen and androgen on the pituitary are very similar. They both increase the number of eosinophil cells,

come from the adrenal cortex. Nevertheless in some cases of the adreno-genital syndrome in women the libido becomes masculine in nature, but this may be due to production of abnormal androgens by the adrenal cortex. Libido and potency, of course, are apt to be fundamentally influenced by psycho-sexual, as opposed to endocrine, impulses and it is difficult in some cases to make a clear differentiation between these two influences. Female libido, for instance, though usually considered to be

and anxious to lead, to make decisions and to be obeyed. The female is gentle, kind, understanding, maternal and submissive. This is not altogether a matter of social habit and tradition. There are males who are too masculine for the moral peace of our present society and they receive judicial penalties, as indeed they should, for their appalling misdemeanours. There are males who are homosexual and males who are transvestites, and it is probably right that their aberrations should come within the orbit of the police, or possibly of the psychiatrist, rather than the endocrinologist. Nevertheless it is a fact that sex hormone therapy has an important influence on these psycho-sexual characteristics. Timid, apprehensive, shy and reclusive eunuchoids are dramatically improved with androgen therapy. They develop initiative, drive and force of character. Aggressive, non-co-operative male criminals imprisoned for sexual offences have become reasonable, disciplined and gentle under treatment with oestrogen (Dunn, 1940). It would be dangerous and indeed contrary to present evidence to plead that sexual offences and crimes are due to endocrine imbalance, but it is surely not unhelpful to suggest that in carefully selected cases, endocrine therapy may be employed in the hope of changing the pattern of the immoral behaviour of these dangerous social misfits.

Androgen has some influence upon mammary development and in certain circumstances may give rise to duct proliferation. This is especially noticeable in the case of adolescent youths who may complain of a hard and tender discoid plaque immediately beneath the nipple—the incorrectly termed 'puberty mastitis'.

blocks ovulation (Sawyer *et al.*, 1949). Acetylcholine applied directly to the gland in the exposed pituitary fossa of the rat leads to pseudo-pregnancy, that is, to secretion of Lt.H (Taubenhaus and Soskin, 1941).

It would appear that the gonadotrophic mechanism is ready for action long before puberty. In the process of autotransplantation of the ovaries into the neck region in immature rats, the blood supply to the ovaries must, of course, be temporarily severed and is not resumed until the graft has established itself in the neck. This transient removal of ovarian inhibition releases pituitary gonadotrophins, which subsequently stimulate the graft, and the transplanted ovaries are found to contain enlarged follicles and mature corpora lutea (Greep and Jones, 1950a).

The Physiological Action of Androgens

It seems that the chief method by which androgens exert their influence is to increase the rate of blood flow to the androgen-sensitive tissues, and this applies particularly to the male accessory sex organs such as the penis, scrotum, vas deferens, epididymis, prostate and seminal vesicles. In this respect they are growth or proliferative hormones, as indeed are the oestrogens when they are operating on their specific accessory sexual organs, such, for instance, as the endometrium, or the uterine muscle in pregnancy. Other specifically male accessory tissues in which this growth or proliferative effect can be recognized are the sweat and sebaceous glands of the skin which become hypertrophied and over-active, so that, for instance, seborrhoea and its offspring acne are typical androgenic characteristics. Another example is the thickening and lengthening of the vocal cords which is a natural accompaniment of puberty in the male, but can be induced, and cannot be reversed in the female by injudicious administration of androgen.

Libido and potency—the ability to initiate and maintain erections and produce emissions—represent another characteristic of the accessory male sexual apparatus. In so far as lack of libido and impotence are associated with complete eunuchoidism, it is evident that these characters are directly due to the secretion of testosterone and not just androgens which may

testosterone propionate decreases the 'alkaline' phosphatase and increases the 'acid' phosphatase, and at the same time leads to an increase in the size of the kidney. There is also a large increase in the arginase content of the kidney (Kochakian, 1944). 17-methylation of the androgen hastens the effect so that methyltestosterone and methylandrostenediol give rise to an immediate increase in arginase activity (Kochakian, 1945).

Perhaps the most interesting property of androgen is its protein-anabolic effect. Testosterone, and indeed the androgenic compounds of the adrenal cortex, lay down protein in muscles, in the bone matrix and in other tissues. There is no doubt that this accounts for the increased muscularity and muscular strength of the average male as compared with the average female. The protein-anabolic effect of androgens, however, depends on the state of gonadal function of the individual. It is most apparent in the castrate and the eunuchoid and unnoticeable in the normal male to whom exogenous androgen may be administered Thus senile osteoporosis which may occur in elderly men is benefited by androgen therapy which lays down protein in the matrix of the bone, and the characteristic muscular asthenia of castrates and eunuchoids is overcome by androgen administration. The debility consequent upon chronic illnesses or serious surgical operations may be due to a negative nitrogen balance which can be reversed by androgen therapy. On the other hand administration of androgen to normal males or to elderly men who show no objective evidence of gonadal deficiency is without endocrine effect though its psychotherapeutic value has been exploited from the time of Brown-Séquard onwards. Increased muscularity, however, is a feature of over-activity of the adrenocortical androgens in conditions of adrenal hyperplasia or tumour. It is probable that this leads to the production of non-physiological androgens

There is a marked excretion of creatine when androgens, especially the 17-methylated compounds such as methyltestosterone, are administered to children with retarded physical

This occurs especially in boys who rapidly acquire a marked degree of virility. Sometimes 'gynaecomastia' may appear in boys with delayed puberty who are being treated with chorionic gonadotrophin to induce androgen secretion, and there is undoubtedly a form of gynaecomastia resulting from relative excess of androgen, as well as the more common form of oestrogen-induced gynaecomastia. Androgens can of course suppress lactation and prevent breast engorgement in women, though not so effectively as oestrogen. They are, however, often capable of relieving the tenderness and swelling of mammary fibroadenosis which becomes especially conspicuous in the premenstrual week.

The effect of androgen on bone is somewhat variable. One effect is to stimulate actual growth of bone tissue, which would result in increased length of the long bones. On the other hand, excessive endogenous production of androgen leads to early closure of the epiphyses which will therefore tend to limit the extent to which the long bones can grow. There is little evidence that either exogenously administered androgen or oestrogen can hasten epiphyseal closure in growing children. The metabolism of bone is very largely in the hands of the sex hormones, for oestrogens stimulate osteoblastic activity and also store calcium, whereas androgens help to lay down protein in the bony matrix and also retain phosphorus and calcium. Cases have been described of severe hypercalcaemia following administration of sex hormones in the treatment of inoperable carcinoma of the breast.

The purely androgenic functions of testosterone take priority when it is administered to castrate experimental animals, but when they have been fulfilled, the effect of the hormone on the kidney next becomes noticeable. It is perhaps especially striking that quite small amounts of the hormone can stimulate the target organ and this suggests that the hormone may act as a catalyst or by affecting certain enzyme systems. For instance, the prostate produces acid phosphatase, and castration leads to great diminution of the serum acid phosphatase level, whereas administration of androgen raises the level. Kochakian and Fox (1944) have studied the enzymes of the kidney and find that

DODDS, E. C., GOLDBERG, L , LAWSON, W. and ROBINSON, R (1938) *Nature*, **141**, 247.

DOISY, E. A , VELER, C. D and THAYER, S. A. (1929). *Amer. J Physiol.* **90**, 329.

DUNN, C. W. (1940). *J. clin Endocrinal*, **1**, 643.

[illegible]

FRAENKEL, L. (1903) *Arch f. Gynäk.* **68**, 438

GIRARD, A. and SANDULESCO, G. (1936). *Helv. chim. Acta.* **19**, 1095.

GIRARD, A , SANDULESCO, G , FRIDENSON, A. and RUTGERS, J J. (1932). *C.R Acad. Sci.* **194**, 909.

[illegible]

p. 350, Wisconsin

GREEP, R. O , VAN DYKE, H. B. and CHOW, B F. (1942). *Endocrinology*, **30**, 635

HALLER, A VON (1766). *Elementa physiologia corporis humani* **7**, 30. Lausanne.

HARRIS, G. W (1948). *Physiol Rev.* **28**, 139.

HARTMANN, M. and WETTSTEIN, A (1934). *Helv Chim Acta*, **17**, 878.

HELLER, C G , PAULSEN, C A , MORTIMORE, G. E , JUNGCK, E C. and NELSON, W O. (1953) *Ann. N Y Acad. Sci.* **55**, 685

HIGHMORE, N (1651). *Corporis humani disquisitio anatomica*, p 90. Hague Comitis

HISAW, F. L (1947) *Physiol Rev* **27**, 95.

HOOKER, C. W and FORBES, T. R (1947). *Endocrinology*, **41**, 158.

HOWARD, R P., SNIFFEN, R. C., SIMMONS, F. A and ALBRIGHT, F (1950). *J clin. Endocrinol.* **10**, 121.

KOBER, S (1931) *Biochem Z* **239**, 209

[illegible]

17

LOEWE, S , VOSS, H E , LANGE, F and WAHNER, A. (1928) *Klin. Wchschr.* **7**, 1376.

MACCORQUODALE, D W , THAYER, S. A and DOISY, E. A (1935) *Proc. Soc exp Biol Med* **32**, 1182

MCCULLAGH, D R (1932). *Science*, **76**, 19

MCCULLAGH, E P , SCHNEIDER, R W , BOWMAN, W. and SMITH, M D (1948) *J clin Endocrinol* **8**, 275

MCGEE, L. (1927) *Proc Instit Med*, *Chicago*, **6**

property of causing retention of sodium, chloride, phosphorus and water. It is a matter of degree and the effect is not so conspicuous with androgens as with deoxycorticosterone or oestrogen, but one must be careful not to overdose elderly men with injudicious androgen therapy lest they develop frank oedema and heart failure.

REFERENCES

ALLEN, E. and DOISY, E. A (1923). *J. Amer. med. Ass.* **81**, 819

ASTWOOD, E. B. (1941). *Endocrinology*, **28**, 309.

BACHMANN, W E., COLE, N and WILDS, A L (1940). *J. Amer. chem Soc.*, **62**, 824

BERTHOLD, A A. (1849). *Arch f. Anat Physiol. u wissensch Med.* p. 42. Berlin

BRADBURY, J. T., BROWN, W. E and GRAY, L. A. (1950). *Recent Progress in Hormone Research*, v, 151. Academic Press, New York.

BROWN-SÉQUARD, C. E. (1889). *Arch de. physiol. norm. et path.* 5s. i. Paris.

BUTENANDT, A. (1929). *Dtsch. med. Wchschr.* **55**, 2171.

192.

5.

siol. Chem. **227**, 84.

CALLOW, N H , CALLOW, R. K. and EMMENS, C. W. (1938) *Biochem. J.* **32**, 1312.

COHEN, A., COOK, J. W. and HEWETT, C L. (1935) *J. chem. Soc.*, Part 1, 445

COHEN, A , COOK, J. W., HEWETT, C. L and GIRARD, A. (1934) *J. chem Soc.*, Part 1, 653

COOK, J. W., DODDS, E C and HEWETT, C. L. (1933). *Nature*, **131**, 56

COOK, J. W. and GIRARD, A. (1934). *Nature*, **133**, 377.

CORNER, G. W. and ALLEN, W M (1928) *Amer J Physiol.* **86**, 74.

DAVID, K , DINGEMANSE, E , FREUD, J. and LAQUEUR, E. (1935). *Z. physiol. Chem* **233**, 281.

DAWSON, A B. (1948) *Anat Rec* **102**, 103

med. Wchschr. **55**, 301.

DODDS, E C, GOLDBERG, L., LAWSON, W. and ROBINSON, R (1938) *Nature*, **141**, 247.

DOISY, E. A., VELER, C. D. and THAYER, S A. (1929) *Amer J. Physiol.* **90**, 329

DUNN, C. W. (1940). *J. clin. Endocrinol* **1**, 643.

EVANS, H M. and SIMPSON, M. E. (1950). *The Hormones*, ed. Pincus, G ii, 351.

EVERETT, J W. (1947). *Endocrinology*, **41**, 364.

EVERETT, J. W. (1948). *Endocrinology*, **43**, 389.

FALLOPIUS, G. (1561). *Observationes anatomicae*. Venet.

[illegible]

Acad Sci **194**, 909.

[illegible]

GREEP, R O, VAN DYKE, H. B and CHOW, B F. (1942). *Endocrinology*, **30**, 635

HALLER, A. VON. (1766). *Elementa physiologia corporis humani* **7**, 30. Lausanne.

HARRIS, G. W (1948). *Physiol Rev.* **28**, 139.

HARTMANN, M. and WETTSTEIN, A. (1934) *Helv. Chim. Acta*, **17**, 878.

HELLER, C. G, PAULSEN, C A., MORTIMORE, G. E, JUNGCK, E C. and NELSON, W. O. (1953) *Ann N.Y. Acad Sci* **55**, 685.

HIGHMORE, N (1651). *Corporis humani disquisitio anatomica*, p. 90 Hague Comitis

HISAW, F. L (1947) *Physiol. Rev.* **27**, 95

HOOKER, C. W and FORBES, T R. (1947) *Endocrinology*, **41**, 158.

HOWARD, R P, SNIFFEN, R. C, SIMMONS, F A and ALBRIGHT, F. (1950). *J clin Endocrinol* **10**, 121.

KOBER, S (1931) *Biochem Z.* **239**, 209

KOCHAKIAN, C D (1944) *J Biol. Chem.* **155**, 579

[illegible]

17

LOEWE, S, VOSS, H E, LANGE, F. and WAHNER, A (1928) *Klin. Wchschr* **7**, 1376

MACCORQUODALE, D W., THAYER, S A. and DOISY, E. A. (1935) *Proc. Soc exp Biol Med.* **32**, 1182.

McCULLAGH, D R (1932). *Science*, **76**, 19

McCULLAGH, E P, SCHNEIDER, R W, BOWMAN, W and SMITH, M D (1948). *J clin. Endocrinol* **8**, 275

McGEE, L (1927) *Proc. Instit Med., Chicago*, **6**

MALPIGHI, M. (1686). *Opera omnia*, t.i. Appendix 'De Ovo Incubato', p 30. London.

MARKEE, J. E., SAWYER, C. H. and HOLLINSHEAD, W. H. (1948). *Recent Progress in Hormone Research*, ii, 117. Academic Press, New York.

[illegible]

[illegible] N, S. J. (1932). *Endocrinology*, **16**, 655.

ROSENHEIM, O. and KING, H. (1932). *Chem. Ind.* **51**, 464, 954.

RUZICKA, L., BRUNGGER, H., EICHENBERGER, E. and MEYER, J. (1934). *Helv Chim. Acta*, **17**, 1389.

SAWYER, C. H, MARKEE, J. E and HOLLINSHEAD, W. H. (1947). *Endocrinology*, **41**, 395.

SAWYER, C. H., MARKEE, J. E and TOWNSEND, B. F. (1949). *Endocrinology*, **44**, 18

SCHWENCK, E. and HILDEBRANDT, F. (1932). *Naturwiss*, **20**, 658.

SEVERINGHAUS, A. E. (1937). *Physiol Rev.* **17**, 556.

SLOTTA, K. H, RUSCHIG, H. and FELS, F. (1934). *Ber.* **67**, 1270.

STOCKARD, C. R. and PAPANICOLAOU, G. N. (1917). *Amer. J. Anat.* **22**, 225.

TAUBENHAUS, M. and SOSKIN, S. (1941). *Endocrinology*, **29**, 958.

VAN DYKE, H. B. *The Physiology and Pharmacology of the Pituitary Body* Vol I (1936). Vol. II (1939) Chicago.

VENNING, E. H. and BROWNE, J. S L. (1936) *Proc. Soc. exp. Biol. Med.* **34**, 792.

WIELAND, H. (1928). Le Prix Nobel, Stockholm.

WILKINS, L., FLEISCHMANN, W. and HOWARD, J. E. (1941). *Johns Hopk. Hosp. Bull.* **59**, 493

WINDAUS, A (1928). Le Prix Nobel, Stockholm.

ZIMMERMANN, W (1935) *Z physiol Chem* **233**, 257.

ZIMMERMANN, W. (1936). *Z. physiol Chem* **245**, 47.

ZONDEK, B. (1930) *Klin. Wchschr.* **9**, 245, 393.

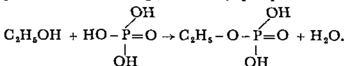
XII

Acid and Alkaline Phosphatase in Disease

E. J. KING

PHOSPHORIC ACID, H_3PO_4 or $HO-P(OH)(OH)=O$, has three replaceable hydrogen atoms. It is therefore capable of forming three series of salts in which one, two, and three hydrogens are replaced by a metal, as for instance mono-sodium phosphate (NaH_2PO_4), disodium phosphate (Na_2HPO_4), and tri-sodium

Phosphoric acid can also form 'organic salts', or esters, in which one or more of its hydrogens is replaced, not by a metallic element, but by a sugar, alcohol, or phenol. Thus, one, two or three molecules of ethyl alcohol can be made to combine with phosphoric acid to form, e.g., mono-ethyl phosphate



$$C_2H_5OH + HO-P(OH)(OH)=O \rightarrow C_2H_5-O-P(OH)(OH)=O + H_2O.$$

Such organic phosphates form a very important class of biological substances, among which are included the sugar phosphates, the phosphatides (lecithin and kephalin), and the nucleotides and nucleic acids. Impressed by the presence of

much phosphorus in nervous tissue, Hegel wrote 'Ohne phosphor, keine Gedank'—'without phosphorus, no thought'. In the light of present knowledge it would be appropriate to extend the metaphor to include *most metabolic processes*, the storage and transfer of energy, indeed life itself. The chemical transformations involved in the formation of these phosphoric esters (synthesis), and in their break-down (hydrolysis), can be brought about in the test tube by the use of strong reagents, *catalysts and heat. In the body their synthesis and hydrolysis are* realized under very much milder conditions by the agency of biological catalysts, called enzymes. The phosphatases are enzymes which synthesize and hydrolyse the 'organic salts', or esters, of phosphoric acid, according to the equation:



$$ROPO_3H_2 + H_2O \rightleftharpoons ROH + H_3PO_4,$$

where ROH is any alcoholic, phenolic, sugar or sugar-like substance, such as ethyl alcohol, phenol, glucose or glycerol. There are several phosphates in both plant and animal tissue, and they are variously active at acid, neutral and alkaline reactions, and are characteristically activated and inhibited by various substances.

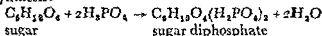
Takahishi (Suzuki, Yoshima and Takahishi, 1907) is usually credited with being the first to describe such an enzyme. From the preparation known as takadiastase, which is made from a fungus growing on bran, he extracted an enzyme which would actively split simple phosphoric esters into the free alcohol or sugar and inorganic phosphate. Early in the century the only esterase known was lipase. Fats are esters of glycerol with fatty acids, e.g. glycero-stearate, and fat-splitting enzymes are ester- [illegible]

[illegible] enzyme, which he showed to be active against glycero-phosphate. Early recognition of similar enzymes, active against several organic phosphates, was made by McCollum and Hart (1908), Neuberg and Karczag (1911), Levene and Medigreceanu (1911), Grosser and Hussler (1912) and Euler and Funke (1912).

Phosphatases in Yeast

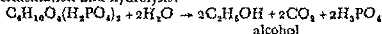
First recognition of the important role of phosphates in living processes came, however, from the work of Harden and Young (1905) on fermentation. The old controversy between Liebig and Pasteur was really never settled until Harden's work became known. Liebig had maintained that alcoholic fermentation was due to an instability, a ferment, set up among the sugar molecules, Pasteur that it was due to the metabolic activity of living organisms (yeasts). Harden showed that active fermentation would take place when a cell-free press-juice of yeast was added to a sugar solution, and that an intermediate step in the production of alcohol was the formation of a sugar-phosphate. The addition of sodium phosphate accelerated carbon dioxide and alcohol production, and a sugar (fructose)-diphosphate—the 'Harden ester'—was isolated from the fermenting solution. This, it appeared, was the essential intermediate substance, which was actively turned into alcohol, carbon-dioxide and free phosphate. The presence in yeast of an enzyme-synthesizing sugar-phosphate was thus demonstrated.

Synthesis:



$$\underset{\text{sugar}}{C_6H_{12}O_6} + 2H_3PO_4 \rightarrow \underset{\text{sugar diphosphate}}{C_6H_{10}O_4(H_2PO_4)_2} + 2H_2O$$

Fermentation and hydrolysis.



$$C_6H_{10}O_4(H_2PO_4)_2 + 2H_2O \rightarrow \underset{\text{alcohol}}{2C_2H_5OH} + 2CO_2 + 2H_3PO_4$$

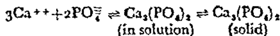
Harden isolated his fructose diphosphoric ester by the addition of calcium or barium, which formed insoluble salts, calcium and barium fructose-diphosphates, which were easily isolated. What Harden did not observe was that there was also formed a sugar mono-phosphate. With his pupil and co-worker, Robison (Harden and Robison, 1914; Robison, 1922), he found that the addition of excess alcohol to the filtrate from the Harden calcium fructose-diphosphate, which Harden had discarded, would yield a copious precipitate of a second phosphoric ester, a sugar mono-phosphate. Its calcium salt was freely soluble in water, but was insoluble in alcohol. This is known as the 'Robison ester'.

In studying the action of yeast on his ester, Robison added some of the clear juice to solutions of his calcium hexose mono-phosphate. The mixture remained entirely clear for some hours, but later he observed heavy chalk-like precipitates at the bottom of the solutions. These were precipitated calcium phosphate, and Robison was not slow to recognize that a phosphoric ester-splitting enzyme must be present in the yeast juice, which hydrolysed his hexose mono-phosphate to free sugar and inorganic phosphate, the latter precipitating as insoluble calcium phosphate. On observing these precipitates and pondering the mechanism of their formation, whereby the soluble calcium hexose phosphate was turned into insoluble calcium inorganic phosphate, Robison was struck by the thought that there might be a similar enzyme present in bony tissue; that, indeed, the formation of bone might be brought about by a process which involved a similar mechanism. He prepared an extract of young bones, by grinding them with water, and found exactly the same result on adding some of it to a solution of his calcium hexose mono-phosphate, as was given by yeast juice. It obviously contained something capable of hydrolysing hexose phosphate, a phosphoric ester-splitting enzyme, a phosphatase. This finding led Robison to his well-known theory of bone formation; and in 1923 he began a series a experiments on this phosphatase enzyme of bone which have led to a clearer conception of the process of bone formation, and a better understanding of the mechanism of calcium phosphate deposition.

The enzyme, which he showed to be present in bone and ossifying cartilage, could be extracted, after breaking up the bone, by treatment with chloroform water. This extract, added to a solution of a primary ester of phosphoric acid, would liberate the phosphorus from its organic combination and, in the presence of inorganic calcium, deposit it as calcium phosphate.

Robison suggested that the bone enzyme in the hypertrophic cells of the tissue where ossification takes place liberates free phosphate from the organic esters of phosphorus contained in the fluids bathing the bone or cartilage, thus giving a local increase in the amount of inorganic phosphorus in solution.

According to the mass law, any increase in the concentration of phosphate-ion would, in the presence of inorganic calcium, lead to a deposition of calcium phosphate.



$$3Ca^{++} + 2PO_4^{\equiv} \rightleftharpoons \underset{\text{(in solution)}}{Ca_3(PO_4)_2} \rightleftharpoons \underset{\text{(solid)}}{Ca_3(PO_4)_2}$$

It was demonstrated that when the bones of rachitic rats were immersed in a solution of calcium glycero-phosphate a dense deposit of calcium phosphate was formed in the matrix of the hypertrophic cartilage, where ossification normally occurs in the animal receiving a proper diet. Since it was only by the action of the bone enzyme that inorganic calcium phosphate could be formed from the glycero-phosphate, both the seat of action of the enzyme and its ability to bring about calcification were thus shown (Plate I, Figure 1).[1]

That there is a natural substrate for this bone phosphatase in the plasma was shown by Martland and Robison (1924, 1926) and Kay and Robison (1924). They showed that the bone phosphatase would act on a trichloroacetic acid extract of blood serum or plasma, which was made slightly alkaline, to liberate free phosphate in a manner similar to its action on a solution of glycero-phosphate. Only a fraction of the total acid soluble phosphorus compounds of the plasma was attacked by the bone phosphatase (at pH 9·0); although King (1932) showed that hydrolysis of the blood phosphoric esters by phosphatase is nearly complete at pH 7 4. There are thus present in *plasma phosphoric esters which form a specific substrate for the* bone phosphatase. Deposition of the solid phase may be brought about in plasma saturated with calcium phosphate by increasing the concentration of phosphate-ion, where the bone enzyme hydrolyses the phosphoric esters contained in the plasma. While the enzyme appeared to be invariably present in bone or ossifying cartilage, unossified cartilage failed to show any phosphatase activity. Working with the bones and cartilages of the human foetus, Martland and Robison failed to find any of the enzyme present in cartilage prior to the appearance of ossification centres, but demonstrated its presence wherever

[1] The plates referred to in this lecture will be found at the end of the book.

ossification was taking place (cf. also King and Hall, 1931; Plate II, Figure 2).

In cultures of embryonic bone (rabbit, chick) Fell and Robison (1929, 1930, 1934) demonstrated the appearance of intense phosphatase activity, about the time that centres of ossification began to appear; and there was a parallelism between the amounts of the enzyme extractable from the cultures and the apparent rates at which calcification was taking place as seen in histological sections. In the dental pulp of embryonic teeth Engel and Furuta (1942) have shown a similar occurrence of the enzyme. On the other hand, the cartilaginous skeletons of the selachian fishes are almost devoid of phosphatase, whereas their teeth are rich in it (Bodansky *et al.*, 1931; Roche and Bullinger, 1939). Huggins (1931) showed that pieces of bladder wall, transplanted intramuscularly, acquired a high phosphatase activity, if and when heterotopic ossification set in. Botterell and King (1935) found an increase of the enzyme in the sites of fracture repair in the radii of rabbits (Figure 3). In addition to these, and many other, chemical demonstrations of the probably important role of phosphatase, many recent histological studies, with new methods of staining for the enzyme, have confirmed the distribution of phosphatase in the skeleton, and its close association with bone formation and many (though not all) processes of calcification (e.g. Bourne, 1943).

Although Kay and Robison (1942) showed the presence of a natural substrate for the bone phosphatase in the blood plasma, the amount of the plasma ester phosphate was, however, very small (Martland and Robison, 1926), and there was some doubt as to whether the quantities of phosphate which its enzymic hydrolysis could liberate would account for all the calcium phosphate deposited in bone. Robison suggested the possibility of a 'phosphate cycle' in the bone, whereby phosphoric esters would be synthesized in one stage of a process, and hydrolysed by the phosphatase in the osteoblasts at a later stage. This idea was worked on by Robison and Rosenheim (1934), and has been given detailed expression by Gutman and Gutman (1941) and by Roche (1947, 1950), who postulate that a phosphoric ester-

synthesizing enzyme (phosphorylase) brings about a phosphorylysis of the glycogen in bone with the inorganic phosphate of the plasma to produce glucose-l-phosphate, which becomes the principal substrate for the bone phosphatase.

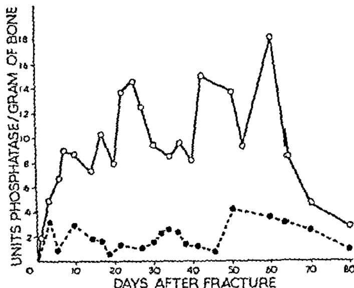


FIG. 3 Increase of phosphatase at the site of fracture in rabbits' radii, shown by continuous line. Radii of opposite arms used as controls shown by broken line.

The Robison theory of bone formation is not universally accepted by all workers, and other roles for the bone phosphatase have been suggested (cf. McKelvie and Mann, 1948); Siffert (1951); Neuman and Neuman (1953) and DiStefano, Neuman and Rouser (1953)). But it serves as a good working hypothesis which explains most, if not all, of the known facts, and there is no doubt that it has stimulated much useful effort and many fruitful ideas.

PLASMA PHOSPHATASE (ALKALINE)

That the blood plasma also contains a phosphatase, similar in properties to that of bone, was shown by Martland and Robison (1926); and the relative phosphatase activities of the plasma

from many clinical conditions were studied by Kay (1929, 1930), who found a marked elevation in generalized bone disease, such as rickets, osteomalacia, Paget's and von Recklinghausen's disease, and normal or nearly normal levels in other diseases.

Properties

pH. The plasma phosphatase, like that of the bone, is markedly dependent on pH. Its optimum stability is between pH 7 and 8. While active at neutrality, and at physiological pH, its activity is much more pronounced at an alkaline reaction. Thus, Robison and Soames (1924) found a very rapid

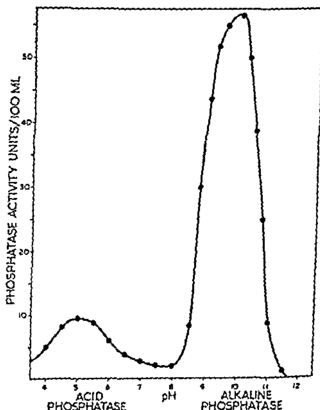


FIG. 4. Variation of plasma phosphatase activity with pH. Plasma from a case of Paget's disease, showing raised alkaline phosphatase (cf. p. 201) and acid phosphatase (cf. p. 206).

increase in the percentage of ester hydrolysed between pH 7 and pH 8·4, but no further change up to 9·4; and at more alkaline pH the hydrolysis rapidly diminished. The 'optimum' pH for maximum hydrolysis by phosphatase varies slightly with the ester-phosphate used as substrate: ethyl-phosphate is optimally hydrolysed at pH 8·1, glucose-phosphate at 8·6, glycero- at 9·4, phenyl- and nitrophenyl- at 9 8 (King and Delory, 1939). As will be seen below, methods of phosphatase determination must take these pH optima into account (Figure 4).

Substrate specificity. The plasma and bone phosphatases (and those of kidney and intestine) catalyse the hydrolysis only of the primary esters, i.e. the mono-esters of orthophosphoric acid. They do not act on secondary (di-) esters. Mono-ethyl phosphate is attacked, but di-ethyl phosphate is not. Nor are they active against the phospho-lipins and other complex organic phosphorus-containing substances of animal tissues.

Activation and inhibition. It is characteristic of the alkaline phosphatase, whether it is from plasma, bone, kidney, or intestine, that its activity diminishes as it is dialysed free of magnesium ions; and that the activity is restored on adding a magnesium salt (Erdtman, 1927). It is thus activated by Mg^{++}. On the other hand it is made less active by the addition of sulphydryl compounds, such as glutathione (Waldschmidt-Leitz *et al.*, 1933), which act as inhibitors. Certain amino-acids activate it, e.g. glycine, alanine (Bodansky, 1946; Roche *et al.*, 1944), histidine (Akamatsu and Kobayashi, 1951). Kutscher and Sieg (1950) have claimed that choline pyrophosphate is its co-enzyeme.

Methods of Estimation

From what has been said under Bone Phosphatase it will be apparent that a measure of the phosphatase activity of a tissue or extract could be gauged from the quantity of calcium phosphate precipitated in a given time, when it was added to a solution of the calcium salt of hexose mono-phosphoric ester. Such crude estimations of phosphatase were not very useful. Although the weights of the precipitates would give a comparative measure of enzyme activity, something much easier and

quicker was required. About the time Robison was commencing work on his enzyme, colorimetric micro methods for inorganic phosphate were being invented, and he took advantage of the blue reduced phospho-molybdate procedure for determining small concentrations of inorganic phosphate to develop a quick and convenient way of determining phosphatase activity. A small measured amount of a standard extract of bone (e.g. 1 part of pulverized bone extracted 48 hours with 50 parts of chloroform water) was added to a standard solution of hexose phosphate, whose organically combined phosphate does not react with molybdate, and the amount of inorganic phosphate liberated by enzymic hydrolysis in a certain time was estimated by adding molybdate and a reducing agent like hydroquinone. The intensity of blue colour produced gave the measure of phosphatase activity in the extract, in terms either of milligrams of phosphate liberated in a given time, e.g. by extract corresponding to 1 gram of bone, or of the percentage of hexose phosphate hydrolysed (Martland and Robison, 1926).

Glycero-phosphate methods. Kay (1930) used the same reaction for his well-known procedure for plasma phosphatase. Instead of hexose phosphate, he used the more easily obtainable sodium β-glycero-phosphate as substrate, and allowed a measured amount of plasma to act on a glycero-phosphate solution at 37° C., and at its own pH of about 7·4. The use of the physiological pH appeared to have the advantage of keeping the hydrolysis at the same reaction as similar hydrolyses were supposed to proceed in the body, but it had the grave disadvantage of making the rate of hydrolysis so slow that only after 48 hours was sufficient inorganic phosphate liberated to give a measurable colour with the molybdate method used. Although Kay's results on cases of bone diseases commanded wide interest, his method of determination took so long that a simpler and quicker procedure was obviously necessary before phosphatase estimations could have any wide clinical application.

Jenner and Kay (1932) modified Kay's method by conducting the hydrolysis at a pH nearer the optimum (pH 8·8, glycine buffer), thus gaining much more analysable inorganic phosphate in a much shorter time. By the use of a very sensitive

method for phosphate (molybdate-stannous chloride) they were able to cut the hydrolysis time to 3 hours, thus producing a serviceable and convenient procedure for clinical laboratory purposes. 0·25 ml. of plasma is allowed to act on the buffered glycero-phosphate for 3 hours, when the hydrolysis is stopped with trichloroacetic acid and the increase of inorganic phosphate (there is a small amount initially present in the plasma) is measured. Their unit of phosphatase activity is defined as that amount of enzyme which will liberate inorganic phosphate corresponding to 1 mg. P from excess glycero-phosphate at 37° C. and pH 8·8 in 3 hours; and the units per 100 ml. of plasma are equal to the number of mg. of inorganic P which are hydrolysed under these conditions. By this method average normal plasmas give about 6 units, with a range of 3 to 13 units. The King and Armstrong method (see below) yields similar values.

Bodansky (1933) has described a rather similar method for phosphatase, which is much used, particularly in America. The hydrolysis is carried out in a barbitone buffer (pH 8·6), and the units are equal to the mg. of P liberated in 1 hour from 0·5 per cent glycero-phosphate, per 100 ml. of plasma or serum. Although it is not possible to equate exactly one unit to another, that of Jenner and Kay is, on the average, about 2½ times the Bodansky unit, e.g. a plasma having 10 Jenner and Kay units would show about 4 units in the Bodansky method.

Shinowara, Jones and Reinhart (1942) and others have described similar methods, based on glycero-phosphate.

Phenyl-phosphate methods. King and Armstrong (1934), in seeking a quicker and more convenient procedure for plasma phosphatase determination, took advantage of the fact that the phosphoric ester of phenol is hydrolysed 2½ times as quickly as glycero-phosphate. Moreover, there was a colorimetric method for phenol (Folin and Ciocalteu, 1927) which was quicker and more sensitive than existing methods for phosphate. Since, in the hydrolysis of phenyl-phosphate, a molecule of phenol is liberated for each molecule of inorganic phosphate, the rate of hydrolysis can be determined as satisfactorily by estimating one as the other. Because of the faster hydrolysis and the sensitive

phenol method, King and Armstrong were able to reduce both the amount of plasma used and the time of hydrolysis. In the original (1934) method the incubation time was 30 minutes and a veronal buffer (pH 9·4) was used. By using the exact pH optimum (pH 9 8 at 37° C.), a more recent modification has reduced the time of hydrolysis to 15 minutes (King, 1945, 1951), thus increasing still further the speed and ease of the determination, as well as its accuracy, since the greatly decreased time of hydrolysis yields results approximating to initial velocities. 0·2 ml. of plasma is mixed with 4 ml. of sodium carbonate - bicarbonate - buffered phenyl - phosphate, kept at 37° C. for 15 minutes, the enzyme action stopped and the proteins precipitated by adding Folin and Ciocalteu's phenol reagent, and the blue colour due to free phenol developed in the filtrate by adding sodium carbonate. The units are equal to the mg. of phenol hydrolysed, per 100 ml. plasma. They are equal to the Jenner and Kay units, and are 2½ times (on the average) the Bodansky units. Several modifications of this method have been described, e.g. by Reiner (1953) and Buch and Buch (1939), Greenberg and Weitzman (1940), and Aebi (1952); and the phenol may be determined by other methods, e.g. diazo (King and Armstrong, 1934; Gomori, 1949; Kaplan and Narahara, 1953) and by amino antipyrene (Grifols-Lucas, 1951; Powell and Smith, 1954; Kind and King, 1954).

p-Nitrophenyl-phosphate methods. Ohmori (1937) was the first to use this substrate for phosphatase estimation; and King and Delory again used it in 1939. It is attacked even faster by the enzyme than is phenyl-phosphate (King and Delory, 1939; Delory and King, 1943). The *p*-nitrophenol which is liberated is of an intense yellow colour at an alkaline reaction and can be determined direct in a photometer or colorimeter without adding any further reagent than strong alkali, which serves both to stop the enzyme action and to bring up the yellow colour of the nitrophenol. Bessey, Lowry and Brock (1946) have based a plasma phosphatase method on this reaction.

Phthalein-phosphate methods. King and Lawford prepared the phosphoric ester of phenolphthalein, which is colourless in alkaline solution because its potential quininoid structure is

blocked by the phosphate group (cf. Lawford, 1937; King, 1938). On hydrolysis by phosphatase, free phenolphthalein is liberated and the solution slowly turns pink. King (1943) described several phthalein and sulphonephthalein phosphates (phenol- and thymol-phthalein, phenol red, thymol blue, bromphenol blue), but thought that none of them made as useful a substrate for phosphatase determination as the phenyl-phosphate of King and Armstrong; although Bray and King (1943) found phenolphthalein-diphosphate to be a useful re-agent for differentiating bacteria. Huggins and Talalay (1945), however, have considered that it furnishes a suitable and convenient method for plasma phosphatase estimation, and have based a procedure and a unitage on the phenolphthalein liberated under defined conditions of pH, temperature, time, etc.

Besides the phenolphthalein-diphosphate, a phenolphthalein-monophosphate has been made by Gutman (1950) and by King (1948, unpublished). It is likewise colourless, and its solution turns pink when hydrolysed by phosphatase in an alkaline buffer. We have used this substance for plasma phosphatase estimations, but regard it also as inferior for the purpose to phenyl-phosphate. Moreover, the several preparations of phenolphthalein-monophosphate which have been made have differed in their rates of phosphatase hydrolysis, and it does not appear that the method of synthesis used yields a single chemical substance.

Plasma phosphatase figures by several methods are given in Table 1.

TABLE 1. Plasma Alkaline Phosphatase by Several Methods (normal ranges)

Kay (1930) units per 100 ml. plasma	0 1–0 2
Jenner and Kay (1932) units per 100 ml plasma	3–13
King-Armstrong (1934) units per 100 ml. plasma	3–13
Bodansky (1933) units per 100 ml. plasma	1·5–5
Shinowara, Jones and Reinhart (1942) units per 100 ml plasma	2 8–8 6
Huggins and Talalay (1945) units per 100 ml plasma	3–15

Normal Values of Alkaline Phosphatase

The two most commonly used methods for plasma phosphatase are the Bodansky and the King-Armstrong. The figures

given for plasma levels in health and disease in the following sections of this paper are in King-Armstrong units. These authors gave the normal range as 3–13 units for adults, and up to 20 units for children (cf. Buch and Buch, 1939). Young, King, Wood and Wootton (1946) studied plasma levels in healthy women in pregnancy in the United Kingdom and Stern (1947) in children and adults in Germany. These representative findings are summarized in Table 2.

TABLE 2. Normal Plasma Alkaline Phosphatase (King-Armstrong units)

Healthy laboratory workers and others	
[illegible]	4·0 – 10·0
[illegible]	(3·3) 4·5 – 9·5 (12·9)
[illegible]	3·0 – 13·0
[illegible]	4·4 – 13·2
[illegible]	2·0 – 8·0
[illegible]	
35 men	4·0 ± 2·0
33 women	3·0 ± 2·1
71 children	12·5 (5–28)
[illegible] (Young et al., 1946)	
[illegible]	4·3 ± 2·0
[illegible]	10·5 ± 3·0
[illegible]	10·4 ± 3·5
Post-partum (6 months)	7·6 ± 2·5

ALKALINE PHOSPHATASE IN BONE DISEASE

Rickets. Many authors have noted an elevation, often marked, of the plasma phosphatase in clinical rickets (Kay, 1929; Bodansky and Jaffe, 1934; Morris and Peden, 1937; Jaffe and

TABLE 3. Gray and Carter's (1949) Results for Plasma Alkaline Phosphatase in Normal and Rachitic Children (King-Armstrong units) Infants, 0–3 years

No.	Diagnosis	Average	Range
56	Normals	17	11–20
64	Clinical rickets	25–40[1]	25–76[1]
	Early rickets	35[1]	

[1] Fall to normal on treatment, or three weeks after administration of 200,000 units of calciferol

Bodansky, 1943; Patwardhan, Chitre and Sukhatankar, 1944; Cantarow and Trumper, 1945). A good summary is given by Gray and Carter (1949), and their own findings (cf. Table 3) are a very useful guide to the use of these estimations in the diagnosis and treatment of this disease.

Osteomalacia. This adult form of rickets also yields enhanced values, which usually are in the 20's of units, and often higher.

Osteitis fibrosa cystica. A high plasma phosphatase is almost invariable in von Recklinghausen's disease. This, with the high serum calcium, is a reliable diagnostic finding. Depending on the state of development of the disease, values of anything up to 200 units (occasionally over) may be found. Usually they are in the 40 to 100 unit range.

Osteitis deformans. Paget's disease likewise has an elevated phosphatase, and of about the same order. But the nearly normal serum calcium serves to give a useful contrast to von Recklinghausen's in which phosphatase and calcium are both high. Paget's disease shows considerable variability in the phosphatase values from week to week, as is shown in Table 4.

TABLE 4 Variation of Alkaline Phosphatase in Paget's Disease
Periodic determinations (King-Armstrong units)

		Case		
1	2	3	4	5
41	52	39	81	33
52	62	54		29
23	36	45	79	37
65	75	45	(4 mth [1])	32
(3 mth [1])	67	66		34
	70	(3 mth [1])		30
	(4 yr [1])			(1 yr.[1])

[1] Period during which investigations were made

Fractures. Although there is a big local increase in the bone phosphatase at the site of a healing fracture (Botterell and King, 1935), this is only mildly reflected in the plasma. Unitages in the 10's are usually found, and occasionally in the 20's (cf. Peden, 1937)

Bone tumours. In general it may be said that any case of bone cancer, in which active replacement is going on, accompanied

by osteoblastic activity, will show a raised plasma phosphatase, e.g. in osteogenic sarcoma, and metastatic carcinoma; but those malignancies of bone which are destructive without replacement, and which show high osteoclastic activity, do not cause excess phosphatase to be made, and no elevation of the enzyme in the plasma occurs (multiple myeloma and Ewing's tumour); cf., for instance, Franseen, Simmons and McLean (1939), Woodard (1942); Sullivan, Gutman and Gutman (1942), Sunderman (1942), King and Delory (1948). A summary of the biochemical findings in bone diseases is given in Table 5.

TABLE 5. Biochemical Findings in Bone Disease
(n = normal; l = low; h = high; sl. = slightly)

	Calcium (mg /100 ml.)	Phosphorus (mg./100 ml.)	Alkaline Phosphatase (units/100 ml.)
Rickets	n (9–11)	l (1–2)	h (25–40)
Osteomalacia	n (9–11)	l (½–1)	h (20–)
Tetany	l (5–7)	n (2–4)	n (3–10)
Osteitis fibrosa	h (12–20)	n (2–4)	h (20–200)
Osteitis deformans	n (9–11)	n (2–4)	h (20–200)
Metastatic carcinoma	n (9–11)	n (2–4)	n (to h) (10–50)
Osteogenic sarcoma	n (9–11)	n (2–4)	h (10–40)
Multiple myeloma	often h (10–)	n (2–4)	n (3–10)
Fractures	n (9–11)	n (2–4)	h (sl.) (10–20)

ALKALINE PHOSPHATASE IN JAUNDICE

Roberts (1933) was the first to show that bone diseases were not unique in having a high plasma phosphatase. He found gross elevation in obstructive jaundice, and nearly normal levels in non-obstructive cases. Armstrong, King and Harris (1934) brought experimental proof of this finding by ligating, or otherwise obstructing, the common bile ducts of dogs, thereby causing enormous increases (up to 500 units) in the plasma phosphatase, as well as in the bile pigments. On relief of the obstruction, the jaundice subsided and the phosphatase slowly re- [illegible] mal levels (Figure 5). Similar results were obtained [illegible] since been confirmed [illegible] Chen and Ivy (1938), Gutman, Hogg and Olson (1940), Dalgaard (1947, 1949), Jack-

son (1952). In contrast to the very high levels encountered in experimental obstructive jaundice, only moderately elevated values were found in toxic jaundice produced by chloroform and phosphorus poisoning of the liver; and experimental haemolytic jaundice (hydrazine, toluidine) caused no rise of the plasma phosphatase (Armstrong and King, 1935).

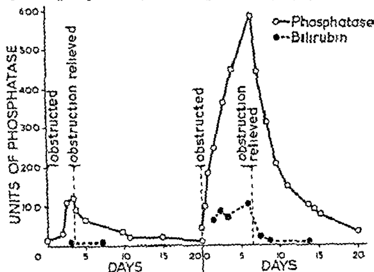


FIG. 5. Obstruction of the common bile duct of the dog. Effect on plasma phosphatase and bilirubin levels (after Armstrong, King and Harris).

Armstrong, King and Harris (1934) described several cases of clinical obstructive jaundice, due, e.g., to carcinoma of the head of the pancreas, in which markedly raised plasma phosphatase values were found (> 100 units); after operation there was a gradual decline until normal levels were reached. Herbert (1935) made an extensive study of the enzyme in the different types of jaundice, as did Maclagan (1944) and Sherlock (1946). Their findings, and our own numerous results, lead to the following summarized conclusions: in obstructive jaundice, in which the obstruction is nearly complete or of long standing, values of over 30 and up to 200 units commonly occur, while less than 30 units may be found if the obstruction is partial or intermittent; in infective hepatitis, cirrhosis and toxic

jaundice about 15–30 units are probable; while the plasmas of haemolytic jaundice cases give phosphatase units within the normal range.

Similar findings, though in terms of different unitages, have been reported by Greene, Shattuck and Kaplowitz (1934), Meranze, Meranze and Rothman (1939), Gutman, Olson, Gutman and Flood (1940), Bodansky and Jaffe (1934), cf. Cantarow and Trumper (1949). Sherlock (1946) gives a valuable table of the findings for plasma phosphatase and other biochemical estimations; and the use that can be made of them in the diagnosis and assessment of liver disease.

Acid Phosphatase

Numerous attempts were made, over several years, to find phosphatase in the urine of normal and diseased subjects; but very little of the enzyme was demonstrated, and no useful variations were found which might be helpful in the investigation of clinical cases. This failure was partly due to the inadequacy of the chemical methods of studying the enzyme in the urine, where the presence of very high concentrations of inorganic phosphate and other substrates interfered strongly with the enzyme hydrolysis of the substances used, and with the estimation of the hydrolysis products. But the principal reason for failing to find anything of interest regarding a phosphatase enzyme in the urine was the emphasis which had always been placed on the alkaline optimum of working of phosphatase. Because people always looked for an enzyme active in the presence of an alkaline buffer, they failed to detect the presence of a potent phosphatase which is active in an acid medium, and present in the urine of males. It is true that Demuth had described an 'acid phosphatase' in 1925, but this seemed to escape the notice of most workers. In 1935, however, two German workers, Kutscher and Wolbergs, detected the presence of a high phosphatase activity with an acid optimum in male urine, and found much less of it in female urine.

Kutscher and Wolbergs reported the presence of large amounts of the acid phosphatase in the prostate (Plate III, Figure 6) and seminal fluid. This phosphatase differed from that

of the bone and kidney by being optimally active in the acid range pH ca. 5. It may therefore be referred to as 'acid phosphatase'. It is present in small amounts in blood plasma, as well as in urine.

Methods of Estimating Acid Phosphatase

The same methods as have been described for alkaline phos- [illegible]

the other reagents and details of the determintion being the same. At this pH the alkaline phosphatase is quite inactive (as the acid phosphatase is at the alkaline pH used for estimating alkaline phosphatase). The phenyl phosphate method of King and Armstrong (1934) is usually used for determining acid phosphatase, but with a citrate buffer of pH 5, and with 1 hour's incubation in order to obtain a measurable amount of released phenol, the amount of acid phosphatase in normal blood plasma being rather small (Gutman and Gutman, 1940; Huggins *et al.*, 1941; Watkinson, Delory, King and Haddow, 1944; King and Delory, 1948). The glycero-phosphate methods may also be used, e.g. that of Shinowara *et al.* (1942), and are preferred by some workers, e.g. Woodard (1942). A procedure based on a diazo reaction has been described by Seligman *et al.* (1951).

Besides the prostatic phosphatase, there are other acid phosphatases in various tissues of the body (liver, spleen, kidney, etc.) (cf. Roche, 1931, Kutscher and Wolbergs, 1953; Roche, Thoai and Baudoin, 1942; Behrendt, 1943; Abul-Fadl and King, 1949). Herbert (1944, 1946) differentiated prostatic phosphatase from that of normal serum by the ready destructibility of the former with alcohol. King, Wood and Delory (1945) found the acid phosphatases of prostate and red cells to be similar in many respects, including easy destruction by ethanol Red blood cells contain roughly 100 times as much acid phosphatase as is present in serum, and haemolysis is, therefore, a potent source of error in the estimation of serum acid phosphatase. Formaldehyde treatment gives a sharp differentiation between the prostatic and red-cell acid phos-

jaundice about 15–30 units are probable; while the plasmas of haemolytic jaundice cases give phosphatase units within the normal range.

Similar findings, though in terms of different unitages, have been reported by Greene, Shattuck and Kaplowitz (1934), Meranze, Meranze and Rothman (1939), Gutman, Olson, Gutman and Flood (1940), Bodansky and Jaffe (1934), cf. Cantarow and Trumper (1949). Sherlock (1946) gives a valuable table of the findings for plasma phosphatase and other biochemical estimations; and the use that can be made of them in the diagnosis and assessment of liver disease.

Acid Phosphatase

Numerous attempts were made, over several years, to find phosphatase in the urine of normal and diseased subjects; but very little of the enzyme was demonstrated, and no useful variations were found which might be helpful in the investigation of clinical cases. This failure was partly due to the inadequacy of the chemical methods of studying the enzyme in the urine, where the presence of very high concentrations of inorganic phosphate and other substrates interfered strongly with the enzyme hydrolysis of the substances used, and with the estimation of the hydrolysis products. But the principal reason for failing to find anything of interest regarding a phosphatase enzyme in the urine was the emphasis which had always been placed on the alkaline optimum of working of phosphatase. Because people always looked for an enzyme active in the presence of an alkaline buffer, they failed to detect the presence of a potent phosphatase which is active in an acid medium, and present in the urine of males. It is true that Demuth had described an 'acid phosphatase' in 1925, but this seemed to escape the notice of most workers. In 1935, however, two German workers, Kutscher and Wolbergs, detected the presence of a *high phosphatase* activity with an acid optimum in male urine, and found much less of it in female urine.

Kutscher and Wolbergs reported the presence of large amounts of the acid phosphatase in the prostate (Plate III, Figure 6) and seminal fluid. This phosphatase differed from that

ml.; but, on the other hand, moderately advanced and early cases gave normal results. They pointed out that an upward trend in values indicated an extension of metastases and that estimation of the serum acid phosphatase from time to time provided an objective method for following the course of the disease.

It has been known for a long time that prepubertal castration in adult life causes regression of the accessory sex glands. Both White in 1893 and Cabot in 1896 reported that castration brought about considerable clinical improvement in men with benign prostatic hypertrophy with decrease in size of the gland.

The thesis was advanced by Huggins, Stevens and Hodges (1941) that malignant prostatic tumour is an overgrowth of adult epithelium cells, and that all types of prostatic epithelium undergo atrophy when androgenic hormones are greatly reduced in amount or inactivated or when oestrogens are injected. Conversely symptoms are aggravated by the injection of androgens. These workers studied the clinical effects of castration in 21 patients with advanced prostatic cancer over a period of 20 months. Four of these patients died, in two unsatisfactory clinical results were obtained, but the remainder improved, as evidenced clinically by increased appetite, progressive gain in weight, improvement in blood picture and lessening of pain.

The findings of Huggins and his colleagues on the effect of administration of oestrogens were confirmed by Kahle, Ogden and Getzoff (1942). Concurrent with clinical changes it was found that both castration and injection of oestrogens lowered the serum acid phosphatase and increased the alkaline phosphatase values. Injection of testosterone had the opposite effect. Watkinson, Delory, King and Haddow (1944) and Daniel, Kind and King (1954) have published findings for plasma phosphatase levels in cases of prostatic carcinoma and some other diseases. The findings with the former group are shown in Table 6 and the latter in Table 7.

Cases of prostatic cancer with radiographic evidence of bone metastases had an increased plasma 'acid' phosphatase, the highest value being 108 units per 100 ml. Since then we have investigated two cases with much higher levels of acid phos-

phatases. The latter is completely destroyed by the inclusion of 0·5 per cent neutral formaldehyde in the buffer-phenylphosphate-serum mixture used for the determination; while the prostatic phosphatase is quite unaffected. Other tissue acid phosphatases, including that of normal serum, are variably inhibited by formaldehyde, but none is so sensitive as the red-cell enzyme (Abul-Fadl and King, 1949).

A high acid phosphatase after formaldehyde treatment (above 5 units) strongly suggests a prostatic origin (values above >3 are 'suspicious'), and in most cases it gives useful information for distinguishing between raised values due to the presence of prostatic phosphatase, and those of different origin. The formaldehyde technique permits the use of haemolysed sera or plasmas which are otherwise unfit for acid phosphatase determination.

Acid Phosphatase in the Diagnosis of Prostatic Carcinoma

Very soon after Kutscher and Wolbergs' observation was made, Gutman and his colleagues (1938) began a series of publications which have yielded much valuable information. They showed that the enzyme is not present until puberty, but may be obtained precociously in the Rhesus monkey by injection of testosterone propionate.

Their most interesting observation was that of the presence of the enzyme in carcinomatous prostatic tissue and at the site of skeletal metastases secondary to carcinoma of the prostate. Pursuing this work, they found that the enzyme is present in the blood serum in a small amount, and that there is an increase of the serum acid phosphatase level in cases of metastatic carcinoma of the prostate gland while patients without metastases gave normal results, that is up to 3 units per 100 ml.

In a series of 130 patients with carcinoma of the prostate and definite X-ray evidence of metastases the Gutmans found that 85 per cent gave a raised serum acid phosphatase, in one instance a value of 1,000 units per 100 ml. being found, but normal results were obtained in the other 15 per cent of cases. Some patients with advanced Paget's disease also showed high acid phosphatase values, that is, greater than 3 units per 100

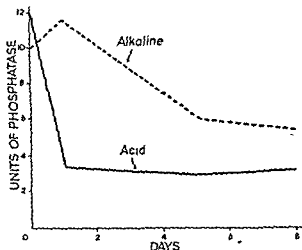


FIG. 7. Plasma acid and alkaline phosphatase in a case of carcinoma of the prostate, with bone secondaries, treated with stilboestrol (after Watkinson *et al.*, 1944).

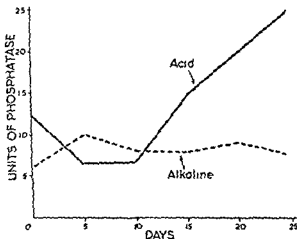


FIG. 8 Rise of plasma acid phosphatase after treatment with testosterone, 25 units daily for 18 days.

phatase, i.e. 426 and 1,700 units, before treatment with stilboestrol. Many of the cases showed regression of the primary tumour under treatment with stilboestrol, and secondary deposits in lymph glands regressed in some cases. In those with bone metastases, three monthly X-ray examinations showed a progressive increase in density in some instances, while in others the deposits became more numerous or increased in size. With one exception, all cases showed some clinical improvement and relief from pain, lessening of frequency of micturition, gradual increase of haemoglobin level, and an improvement in appetite and gain in weight. The changes in plasma phosphatase content during treatment in a typical case are shown in Figure 7. Testosterone has the reverse effect of stilboestrol (Figure 8).

TABLE 6. Acid Phosphatase in Prostate and Blood

	Units per 100 g. or ml.	
	Range	*Average*
Prostatic tissue	50,000–250,000	—
Seminal fluid (12 normals)	87,000–330,000	206,000
(2 eunuchs)	2,000–2,400	—
Blood plasma:		
Males, normal	0–4 6	2 2
Carcinoma of prostate (114 cases)	5·3–108	22 4
(Alkaline phosphatase)	(7–32)	(15·7)
Females, normal	0–4	2
Carcinoma of breast	3 8–6 3	4 9
(Alkaline phosphatase)	(13–22)	(19)

Hock and Tessier's (1949) observation that massage of non-malignant prostates may raise the serum acid phosphatase suggested that ordinary palpation might have a similar effect (Daniel and Van Zyl, 1952). In three of 34 patients (60–80 yr.) receiving routine palpation of their prostates, we found raised serum acid phosphatase levels, suggesting the presence of carcinoma of the prostate. Of 19 prostates examined histologically 14 showed benign hypertrophy, two carcinoma and three were normal (Table 7). The three patients with raised serum acid phosphatase had cystic benign hypertrophy of their prostates. Palpation may have ruptured the walled cysts, releas-

the urines collected in three parts suggested that the first part of micturition washes out the prostatic secretion that has accumulated in the urethra, and that straining, or some other factor associated with the completion of micturition, produces an immediate discharge of more prostatic secretion. The volume

TABLE 8. Average Urinary Excretion of Acid Phosphatase
(From Daniel, Kind and King, 1954)

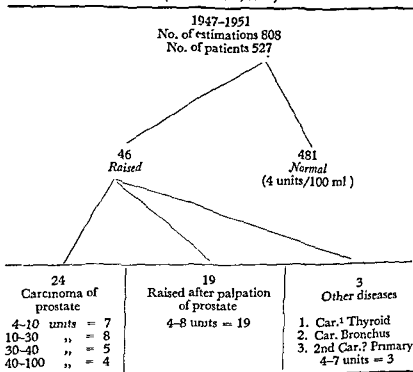
	Total excretion of urinary acid phosphatase (units/ 24 hr.)	Renal excretion of acid phosphatase (units/ 24 hr.)	Prostatic excretion of acid phosphatase (units/ 24 hr.)	Serum acid phosphatase (units/ 100 ml.)
Apparently healthy men 24–35 yr.	1,462	1,074	377	1·2 (0·8–2·1)
Patients with benign prostatic hypertrophy	262	234	45	1·5 (0·4–2·6)
Patients with carcinoma of prostate		128[1]		37·5 (2–144)
4 with impaired renal function		85[1]		
3 with normal renal function		212[1]		
Apparently healthy women 23–60 yr.	217	217		0·75 (0·5–1)
Women with breast carcinoma	132	132		2·0 (1·4–3·1)
Women with breast carcinoma and bone secondaries	227	227		2·5 (1–4·5)

[1] Catheter.

of prostatic secretion required to produce these effects is probably small, as the seminal plasma is rich in phosphatase—for example, that of one of the subjects contained 6,000 units o acid phosphatase per ml. (Plate IV, Figure 9). The assumption that mid-stream urine is reasonably free of prostatic secretion was supported by a close correspondence between the acid phosphatase content of catheter and mid-stream specimens, although an occasional occurrence of unusually high acid phos-

ing phosphatase-rich secretion into the blood. If the serum acid phosphatase, performed within 24 hours of palpation, is raised, it should be repeated a few days later. Elevation of the phosphatase only after palpation or massage suggests cystic benign *prostatic hypertrophy*, and *should* not be regarded as evidence of a carcinoma of the prostate.

TABLE 7. Serum Acid Phosphatase Estimations on Male Patients (From Daniel, 1952)


1947–1951
No. of estimations 808
No. of patients 527

46 *Raised*

481 *Normal* (4 units/100 ml)

24 Carcinoma of prostate	19 Raised after palpation of prostate	3 Other diseases
4–10 *units* = 7 10–30 ,, = 8 30–40 ,, = 5 40–100 ,, = 4	4–8 units = 19	1. Car.[1] Thyroid 2. Car. Bronchus 3. 2nd Car.? Primary 4–7 units = 3

[1] Car = carcinoma

URINARY EXCRETION OF ACID PHOSPHATASE

Daniel, Kind and King (1954) showed that healthy young adults excreted a considerable amount of acid phosphatase in their urines. Of this between one-third and one-fifth could be attributed to the direct addition of prostatic secretion to the urinary stream. Comparison of the acid phosphatase content of

of this enzyme might possibly have given an earlier clue to the diagnosis of carcinoma of the prostate than plasma estimations alone, but the results appear to indicate that this is not the case. The simple determination of formol-stable plasma acid phosphatase therefore remains the best laboratory aid to the diagnosis of prostatic carcinoma.

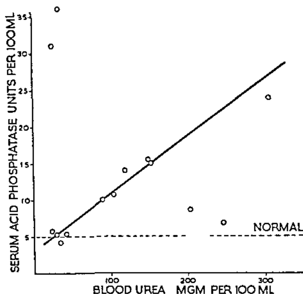


FIG 10. Serum acid phosphatase and blood urea in cases of carcinoma of the prostate.

SUMMARY

The history of the discovery of the phosphatases, enzymes which hydrolyse esters of phosphoric acid, has been reviewed; Robison's phosphatase theory of bone formation is described; methods for the determination of phosphatase in blood plasma are outlined; and the investigation and the use of alkaline and acid phosphatase levels in disease are reviewed.

Plasma alkaline phosphatase is present to the extent of 3–13 units (King-Armstrong) in normal adults, and 11–20 units in normal children. Raised values are encountered in generalized bone disease, e.g. rickets, osteomalacia, osteitis fibrosa cystica,

phatase levels in mid-stream urine could probably be attributed to added prostatic secretion.

Men of 50 and over, with benign prostatic hypertrophy, excreted approximately one-sixth the urinary acid phosphatase of younger men, and this diminution applied to both the renal and the prostatic excretion of acid phosphatase. The output of acid phosphatase in the urine of old men corresponded closely to that of women aged 20 to 60 years. Men suffering from carcinoma of the prostate excreted slightly less urinary acid phosphatase than men of a similar age suffering from benign prostatic hypertrophy; and the difference was most pronounced when renal function was poor. There was no evidence that the excretion of urinary acid phosphatase was related in any way to serum acid phosphatase levels. The acid phosphatase put into the serum by carcinoma of the prostate does not appear to be eliminated in an active way by the kidneys.

Women suffering from carcinoma of the breast do not excrete an increased amount of acid phosphatase in their urine, even though the plasma level may be increased. The fact that women as well as men excreted a considerable amount of formol-stable acid phosphatase in their urine (217 units per day) shows that there must be some source of such phosphatase other than the prostate. These results are summarized in Table 8.

Plasma Acid Phosphatase and Kidney Function

Daniel (1952) compared the plasma acid phosphatase and blood urea of patients with carcinoma of the prostate, and showed, with some exceptions, a rough linear relationship (Figure 10). But uraemia alone did not produce a raised plasma acid phosphatase and none of 20 patients with benign prostatic hypertrophy, who had blood ureas of 60 to 300 mg. per 100 ml., had a persistently raised plasma acid phosphatase.

A persistently raised serum acid phosphatase, especially if over 10 units per 100 ml., is very strong evidence of carcinoma of the prostate, but there is a pressing need for a test which diagnoses the disease in its earliest stages. Had the acid phosphatase produced by carcinoma of the prostate been excreted mainly in the urine, estimation of the rate of urinary excretion

ARMSTRONG, A. R., KING, E. J. and HARRIS, R. I. (1934). *Canad med. Ass J.* **31**, 14.
ARNER, O. and SWEDIN, B. (1949). *Acta chir. scand* 97, 137.
BEHRENDT, H. (1943) *Proc. Soc. exp Biol, N.Y.* **54**, 268.
BESSEY, O. A., LOWRY, O. H and BROCK, M. J. (1946). *J. biol. Chem.* **164**, 321.

31, 1179.

iol. Chem **94**, 551.
BOTTERELL, E. H. and KING, E J. (1935). *Lancet*, **1**, 1267.
BOURNE, G. (1943). *Quart. J. exp Physiol.* **32**, 1.
BRAY, J. and KING, E J. (1943). *J. Path. Bact.* **55**, 315.
BUCH, I. and BUCH, H (1939) *Acta med Scand.* **101**, 211.
CABOT, A. T. (1896). *Ann. Surg.* **24**, 265.
CANTAROW, A. and TRUMPER, M. (1945). *Clinical Biochemistry*, 3rd ed Saunders, London.
CANTAROW, A. and TRUMPER, M (1949). *Clinical Biochemistry*, 4th ed. Saunders, London.

DANIEL, O and VAN ZYL, J J (1952) *Lancet*, **i**, 998.
DELORY, G. E and KING, E J (1943). *Biochem J.* **37**, 547.
DEMUTH, F (1925) *Biochem Z.* **1599**, 415.
DISTEFANO, V., NEUMAN, W. F. and ROUSER, G. (1953). *Arch Biochem. Biophys* **47**, 218
ENGEL, M. B and FURUTA, W (1942) *Proc. Soc. exp. Biol, N.Y.* **50**, 5
ERDTMAN, H (1927). *Hoppe-Seyl Z.* **172**, 182
EULER, H VON, and FUNKE, Y. (1912). *Hoppe-Seyl Z* **77**, 488.
FELL, H. B and ROBISON, R (1929). *Biochem. J.* **23**, 767.

939) *Surg. Gynec Obstet* **68**, 1038.
FREEMAN, S, CHEN, Y P and IVY, A C (1938) *J biol Chem* **124**, 79
GOMORI, G (1949) *J. Lab clin Med* **34**, 275

led. **25**, 634.
J. clin. Invest.
14, 1079

osteitis deformans, certain bone tumours and (slightly) in healing fractures.

In obstructive jaundice high or very high values are obtained if the obstruction is nearly complete or prolonged. Alkaline phosphatase is moderately raised in incomplete or intermittent obstructive jaundice, in hepatitis and in cirrhosis, and is normal in haemolytic jaundice.

The acid phosphatase of the plasma consists of two fractions, prostatic and non-prostatic (partly or mainly derived from the red cells). The prostatic phosphatase is given by the formol-stable acid phosphatase determination. Normal levels are 0–5 units, with the majority below 3. The acid phosphatase is increased in the plasma of cases of prostatic carcinoma with metastases of the bone, when values above 5 and as high as 1,000 units may be encountered, the more usual range, however, being from 5 to about 50. The clinical improvement which results from treatment with oestrogens is accompanied by a decrease of the plasma acid phosphatase. In patients with benign prostatic hypertrophy normal values occur.

Normal healthy men excrete large amounts of acid phosphatase in their urine. Part of this appears to come directly from prostatic secretion, and part from renal excretion. The urinary acid phosphatase of patients with carcinoma of the prostate, and with benign prostatic hypertrophy, is much less than that of normal men. Healthy women, and women suffering from carcinoma of the breast, likewise excrete acid phosphatase in their urine, values being about one-fifth or one-sixth of those of normal males. The source of this female prostate-like acid phosphatase is not known.

REFERENCES

[illegible] *Path* 1 80.
[illegible] 51.
[illegible]
[illegible]
[illegible] 52, 379.

LAWFORD, F. H. (1937). *The enzymatic hydrolysis of phosphoric esters.* Thesis for Ph D. degree of University of Toronto.
LEVENE, P. A. and MEDIGRECEANU, F. (1911). *J. biol. Chem.* **9**, 65
MCCOLLUM, E. V. and HART, E. B. (1908). *J. biol. Chem.* **4**, 497.
MCKELVIE, A. M. and MANN, F. C. (1948). *Proc. Mayo Clin.* **23**, 449.
MACLAGAN, N. F. (1944). *Brit. med J.* **2**, 363.
MARTLAND, M. and ROBISON, R. (1924). *Biochem. J.* **18**, 1354.
MARTLAND, M. and ROBISON, R. (1926) *Biochem. J.* **20**, 847.
MERANZE, T., MERANZE, D. R. and ROTHMAN, M. M. (1939). *Rev. Gastroenterol*, **6**, 254.
MORRIS, N. and PEDEN, O D. (1937). *Quart. J Med.* **6**, 211.
NEUBERG, C. and KARCZAG, L. (1911). *Biochem. Z* **36**, 60.
NEUMAN, W F. and NEUMAN, M. (1953) *Chem. Rev* **53**, 1
OHMORI, Y. (1937) *Enzymologia*, **4**, 217.
PATWARDHAN, V. N, CHITRE, R. G. and SUKHATANKAR, D. R. (1944). *Indian J med Res.* **32**, 31.
PEARSE, A. G. E. (1953). *Histochemistry* Churchill, London
PEDEN, O D (1937). *Arch Dis Childh.* **12**, 87
PLIMMER, R H. A. (1913) *Biochem J.* **7**, 43.
POWELL, M. E. A. and SMITH, M J. H. (1954). *J clin. Path.* **7**, 245.
REINER, MIRIAM (1953). *Standard Methods of Clinical Pathology* Acad. Press, New York.
ROBERTS, W. M. (1933). *Brit med J* **1**, 734.
ROBISON, R (1922) *Biochem. J* **16**, 809.
ROBISON, R and ROSENHEIM, A. H (1934). *Biochem. J* **28**, 684.
ROBISON, R and SOAMES, K. M. (1924) *Biochem J* **18**, 740.
ROCHE, J (1931) *Biochem J* **25**, 1724
ROCHE, J (1947). *Ann. Nutr. Paris*, **1**, 3.
ROCHE, J (1950). In *The enzymes chemistry and mechanism of action*, Sumner, J B. and Myrback, K. D R ed., vol 1, pt. 1, p 503. Acad. Press, New York
ROCHE, J. and BULLINGER, E. (1939). *Bull. Soc Chim. biol*, *Paris*, **21**, 166.
ROCHE, J, THOAI, N-V and BAUDOIN, J (1942). *Bull. Soc Chim. biol*, *Paris*, **24**, 247
ROCHE, J, THOAI, N-V and ROGER, M. (1944). *Bull. Soc Chim biol.*, *Paris*, **26**, 1047
SELIGMAN, A M, CHAUNCEY, H H, NACHLAS, M M., MANHEIMER, L H. and RAVIN, H A (1951) *J biol Chem* **190**, 7.
SHERLOCK, S P V (1946) *J Path Bact.* **58**, 523.
SHINOWARA, G Y, JONES, L M and REINHART, H. L. (1942). *J biol. Chem.* **142**, 921
SIFFERT, R S (1951) *J exp Med.* **93**, 415
STERN, M I (1948) *Brit J Nutr.* **1**, 182
SULLIVAN, T J, GUTMAN, E. B and GUTMAN, A. B. (1942). *J Urol., Balt.* **48**, 426.

GRIFOLS-LUCAS, J. A. (1951). *Int. Cong. clin. Path.* London.
GROSSER, P. and HUSSLER, J. (1912). *Biochem. Z.* **39**, 1.
GUTMAN, A. B (1950). *Fed. Proc.* **9**, 180.
GUTMAN, A. B. and GUTMAN, E. B. (1941). *Proc. Soc. exp. Biol., N.Y.* **48**, 687.
GUTMAN, A. B., HOGG, B. M. and OLSON, K. B. (1940). *Proc Soc. exp. Biol, N.Y.* **44**, 613.
GUTMAN, A. B., OLSON, K. B., GUTMAN, E. B. and FLOOD, C. A. (1940). *J. clin. Invest.* **19**, 129.
GUTMAN, E. B. and GUTMAN, A. B. (1940). *J. biol. Chem.* **136**, 201.
HARDEN, A. and ROBISON, R. (1914). *Proc. Chem Soc.* **30**, 16.
HARDEN, A. and YOUNG, W. J. (1905). *Proc. Chem. Soc., Lond* **21**, 189.
HARDEN, A. and YOUNG, W. J. (1908). *Proc. roy. Soc., B.* **80**, 299.
HERBERT, F K. (1935). *Brit. J. exp. Path.* **16**, 365.
HERBERT, F. K. (1944). *Biochem J.* **38**, xxiii.
HERBERT, F. K. (1946). *Quart. J. Med.* **15**, 221.
HOCK, E. and TESSIER, R. N. (1949). *J. Urol., Balt.* **62**, 488.
HUGGINS, C B. (1931). *Biochem. J.* **25**, 728.
HUGGINS, C B, STEVENS, R. E. and HODGES, C. V. (1941). *Arch. Surg.* **43**, 209.
[illegible]
831.
[illegible]
KAHLE, P. J., OGDEN, H. D. and GETZOFF, P. L. (1942). *J. Urol., Balt.* **48**, 83.
[illegible] *clin. Med.* **41**, 819.
[illegible]
KAY, H. D. and ROBISON, R. (1924) *Biochem. J* **18**, 755.
KIND, P. R N and KING, E J. (1954). *J. clin. Path* **7**.
KING, E. J. (1932) *Biochem J* **26**, 1697.
KING, E. J (1938). *J. Soc. chem. Ind., Lond.* **57**, 85
[illegible]
lon.
hill, London.
KING, E. J. and ARMSTRONG, A R (1934). *Canad med. Ass. J* **31**, 376.
KING, E. J. and DELORY, G. E. (1938) *Biochem J* **32**, 1157.
KING, E. J. and DELORY, G. E (1939). *Biochem J* **33**, 1185
KING, E. J. and DELORY, G. E (1948) *Postgrad med J.* **24**, 299
KING, E. J. and HALL, G. E (1931). *Canad med Ass. J.* **25**, 44

XIII

Body Water Control

G. M. BULL

THE total water content of the body is maintained within very narrow limits, but the mechanisms responsible for this constancy are poorly understood. I propose to review some of the relevant information, and to present certain speculations and investigations into the problem of volume-control of body water.

THE CONTROL OF THE RELATIVE VOLUMES OF FLUID IN THE BODY COMPARTMENTS

It is almost certain that the *relative* volumes of fluid in the three main water compartments of the body (cellular, extracellular and vascular) are governed by comparatively simple physico-chemical processes of diffusion of water along osmotic gradients. For instance, movement of water between vascular and extra-cellular compartments is largely governed by the osmotic pressure of plasma proteins, and by intravascular pressure as suggested by Starling many years ago (38). Squire (36), in his recent study of nephrosis, pointed out that Starling's hypothesis, with minor modifications, accounted for the known facts.

Movements of fluid between cellular and extracellular compartments take place when the ionic concentration in one compartment is altered. For instance, if the extracellular ionic concentration is raised, water moves from the cellular compartment to the extracellular and vice versa (11). Other factors such as pH and oxygen tension changes (28) also cause movements of fluid into and out of cells.

SUNDERMAN, F. W. (1942). *Amer. J. clin. Path.* **12**, 404.
SUZUKI, U. YOSHIMA, K. and TAKAHISHI, M. (1907). *Bull. Coll. Agric. Tokyo*, **7**, 503.
WALDSCHMIDT-LEITZ, E., SCHARIKOVA, A. and SCHAFFNER, A. (1933) *Hoppe-Seyl. Z.* **214**, 75.
WATKINSON, J. M., DELORY, G. E., KING, E. J. and HADDOW, A. (1944). *Brit. med. J.* **2**, 492.
WHITE, J. W. (1893). *Brit. med. J.* **2**, 575.
WOODARD, H. Q. (1942). *Cancer Res.* **2**, 497.
YOUNG, J., KING, E. J., WOOD, E. and WOOTTON, I. D. P. (1946). *J. Obstet Gynaec. Brit. Emp.* **53**, 251.

The effector system

It is clear that the main regulation of body water volume takes place through the kidney's increasing or decreasing the volume of urine passed. It is, furthermore, almost certain that in addition to the excretion of water, it must regulate the excretion of electrolytes. There is, in fact, much to suggest that changes in electrolyte excretion, and particularly Na excretion, are primary. For example, the feeding of extra water to a patient with cardiac failure does not usually lead to weight gain, but the feeding of even small quantities of Na is followed by further fluid retention and gain in weight (42).

The details of the renal excretion of water and salt have received a great deal of attention, and there has long been hot debate as to whether changes in glomerular filtration or tubular reabsorption are most important in altering their excretion. This problem is probably insoluble, but in any event, I think of little importance.

The efferent pathway

The efferent pathway is probably largely humoral, because denervation of the kidney (20), or even its transplantation to an abnormal site, does not abolish its ability to regulate body water content (13). Furthermore, in a number of clinical states of excess fluid accumulation, such as congestive cardiac failure, splanchnic block does not alter water or salt excretion (24). The nature of the humoral factors concerned is uncertain. Adrenalin or noradrenalin may be important (35) and, surprisingly, regulation is possible even in the absence of adrenal cortex (29).

The receptor

Even less is known of the receptor. We do not know what it is sensitive to, but I think we have certain clues. In almost all physiological circumstances and clinical states in which the rate of excretion of water *and* salt is altered, there is also an alteration in the circulation. It seems at least likely that in some way an alteration in the circulation plays an important role in providing the stimulus to the receptor.

Control of Concentration

Verney (39) has demonstrated the existence of an extremely sensitive mechanism controlling the osmotic pressure of extracellular fluid. When the extracellular molar concentration is raised, without a parallel rise in intracellular molar concentration, a receptor, which he has called the osmoreceptor, causes the pituitary to release antidiuretic hormone, which in turn increases the tubular reabsorption of water in the kidneys. As a result of the retention of water in excess of solids, the concentration of the extracellular fluid returns towards normal. When the extracellular fluid molar concentration falls, less ADH is released and a greater quantity of water passed, again correcting the disturbance.

The Need to Postulate a Further Receptor System

The osmoreceptor system, by itself, could not regulate the total volume of fluid in the body, because it is only sensitive to changes in osmotic pressure. If excess of isotonic fluid were to accumulate in the body, it could not stimulate the osmoreceptor. It would distribute itself in the various compartments by diffusion, and the relative amount in each compartment would be determined by the solutes accompanying it. For instance, isotonic saline would accumulate in the extracellular and vascular compartments, while an isotonic potassium solution would accumulate in the cellular compartment. As Borst points out, if the regulating systems we have considered thus far were the only ones, we would be in a very sorry plight. We would just swell. It is necessary to postulate a further regulating system or systems to control the absolute volume of fluid in the body. This was pointed out six years ago by Peters (27) and Borst (4), but until very recently the matter has received scant attention. Let us call this postulated regulating system the volume regulator, and examine its likely mode of action.

The Volume Regulator System

There must be at least three components, a receptor, an efferent pathway and an effector system. Let us deal with them in the reverse order.

Baroreceptors

There are so many possible sites at which a pressure change might act, that I do not propose to consider very many.

(*a*) *Change in central venous pressure* is very unlikely to provide the signal. The same reaction of circulatory inadequacy occurs in both oligaemic shock and congestive cardiac failure—conditions in which the venous pressure change is in opposite directions. Furthermore, even in congestive heart failure there is no good correlation between fluid retention and venous pressure (23).

(*b*) *Changes in general arterial pressure* are also unlikely to provide the signal, because in mild oligaemic shock or with light venous

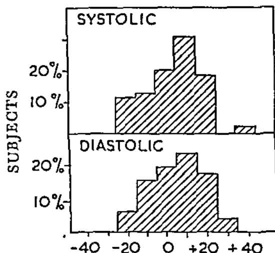


Fig 1. Change in blood pressure on changing posture from lying to standing. The pressures were measured by the auscultatory method, with the arm at approximately the level of the auricles. Fifty normal subjects aged 15 years examined (from Bull, unpublished).

cuffing in the recumbent subject obvious water and salt retention may occur without B P. fall (8). If a normal subject, who has been lying down, stands up, his blood pressure may rise or fall, or remain the same, but he always has a fall in urine volume and salt excretion (8). (Figure 1.)

(*c*) *Intracranial site.* There is some evidence that the receptor we are seeking is intracranial, or has an intracranial component.

Examples of this association between altered circulation and water and salt excretion include the following:

(*a*) *Posture change.* On changing the posture from recumbent to erect the following changes occur: A reduction in excretion of water and salt, and of renal blood flow (7, 15); a fall in venous pressure and cardiac output (26); a fall in blood flow through muscle (6), splanchnic area (10) and brain (31).

(*b*) *Haemorrhage or the application of venous occluding cuffs to the limbs of recumbent subjects* produces the same type of change (41, 17, 40).

(*c*) *Congestive cardiac failure* is associated with the same alterations in excretion of water and salt, and of renal blood flow as above (34), and there is also a reduced blood flow through muscle (6), splanchnic area (25) and brain (30), but the cardiac output may be high or low, and the venous pressure is usually high (12).

(*d*) *Expansion of blood volume* by transfusion of plasma (44), or even saline solutions (21), and exposure to cold (1) produce the reverse change to that described in (*a*). Cold probably causes an increase in effective circulating blood volume by diminishing the peripheral vascular bed, so diverting more blood into the central circulation.

Thus, there appear to be reaction patterns to circulatory inadequacy (*a*, *b* and *c*), and circulatory excess (*d*). In fact, as far as I know, it is impossible to produce an alteration in the general circulation without, at the same time, altering the excretion of water and salt.

Therefore, a possible approach to the problem of the receptor is to consider ways in which altered circulation could affect a receptor. It seems to me that there are only two:

1. By causing a change in pressure in a body compartment or subcompartment which affects a baroreceptor (probably a stretch receptor).

2. By causing a chemical change affecting a *chemoreceptor.* The change in circulation could alter the rate of delivery of a chemical substance to the tissues, or the rate of removal of a waste product.

BARORECEPTORS

There are so many possible sites at which a pressure change might act, that I do not propose to consider very many.

(*a*) *Change in central venous pressure* is very unlikely to provide the signal. The same reaction of circulatory inadequacy occurs in both oligaemic shock and congestive cardiac failure—conditions in which the venous pressure change is in opposite directions. Furthermore, even in congestive heart failure there is no good correlation between fluid retention and venous pressure (23).

(*b*) *Changes in general arterial pressure* are also unlikely to provide the signal, because in mild oligaemic shock or with light venous

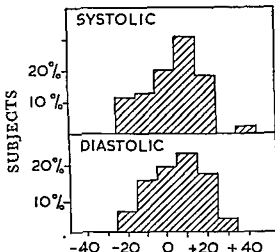


FIG. 1. Change in blood pressure on changing posture from lying to standing. The pressures were measured by the auscultatory method, with the arm at approximately the level of the auricles. Fifty normal subjects aged 15 years examined (from Bull, unpublished).

cuffing in the recumbent subject obvious water and salt retention may occur without B.P. fall (8). If a normal subject, who has been lying down, stands up, his blood pressure may rise or fall, or remain the same, but he always has a fall in urine volume and salt excretion (8). (Figure 1.)

(*c*) *Intracranial site.* There is some evidence that the receptor we are seeking is intracranial, or has an intracranial component.

Welt described a salt-losing state in association with certain

intracranial component, but it is at least worth investigating the possibility that there is an intracranial baroreceptor. Lewis *et al.* (22) believe that they have evidence for the existence of such a receptor. In the normal subject a change of position from the recumbent to the sitting is associated with a fall in urine volume and excretion of salt. Lewis *et al.* state that this can be prevented by the application of a venous occluding cuff to the neck, suggesting that there is a baroreceptor in the head. Their work has not been confirmed. It is necessary in studies on this problem to include most rigorous controls, and I believe that these were inadequate in their study. Barbour and I attempted to confirm their finding, using what we believe to be more satisfactory controls, and found no evidence of altered water and salt excretion on applying a venous occluding cuff to the neck (Figure 2).

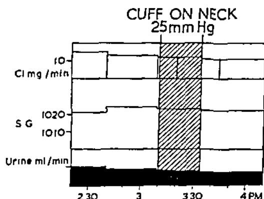


Fig. 2. The absence of effect of applying a venous occluding cuff to the neck of a sitting normal subject on urinary excretion (from Bull, unpublished).

We also attempted to show the presence of an intracranial baroreceptor in another way. We lowered CSF pressure by ... and followed the excretion of water and salt.

It is, therefore, unlikely that there is a baroreceptor in a low-pressure site in the head. Some other intracranial receptor is, however, not excluded.

(*d*) *Intrathoracic site.* An intrathoracic site is another possibility, and indeed I believe there is good evidence for the existence of a receptor which is capable of body water regulation. Sieker, Gauer and Henry (32, 33), and Drury, Henry and Goodman (14) investigated the effects of breathing against positive and negative pressure. They argued that the change in intrathoracic pressure resulting from these manœuvres would change the volume of blood in the low-pressure sites in the thorax. They found that breathing against a low pressure caused a water diuresis, and breathing against positive pressure an antidiuresis. These changes would be of the type expected in a regulator of body water. Thus, excess fluid retention would raise the volume of fluid in the pulmonary veins, and so lead to a diuresis. Diagrammatically, this system would regulate body water as follows (Figure 3):

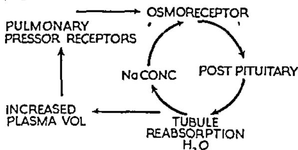


Fig. 3 Diagrammatic representation of a possible method of body water control.

However, this system cannot be the one we are seeking. Neither negative nor positive pressure-breathing affect sodium excretion, an invariable accompaniment of altered circulation.

Borst's Hypothesis

Borst has put forward an hypothesis as follows (4):

The compartment, whose volume is primarily controlled, is the vascular compartment.

Welt described a salt-losing state in association with certain

intracranial component, but it is at least worth investigating the possibility that there is an intracranial baroreceptor. Lewis *et al.* (22) believe that they have evidence for the existence of such a receptor. In the normal subject a change of position from the recumbent to the sitting is associated with a fall in urine volume and excretion of salt. Lewis *et al.* state that this can be prevented by the application of a venous occluding cuff to the neck, suggesting that there is a baroreceptor in the head. Their work has not been confirmed. It is necessary in studies on this problem to include most rigorous controls, and I believe that these were inadequate in their study. Barbour and I attempted to confirm their finding, using what we believe to be more satisfactory controls, and found no evidence of altered water and salt excretion on applying a venous occluding cuff to the neck (Figure 2).

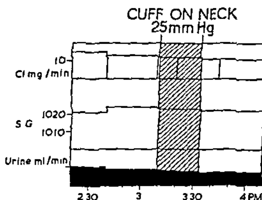


FIG. 2. The absence of effect of applying a venous occluding cuff to the neck of a sitting normal subject on urinary excretion (from Bull, unpublished).

We also attempted to show the presence of an intracranial baroreceptor in another way. We lowered CSF pressure by lumbar puncture, and followed the excretion of water and salt. Within certain limits we found no change (2). This was also the finding of Fishman (16).

It is, therefore, unlikely that there is a baroreceptor in a low-pressure site in the head. Some other intracranial receptor is, however, not excluded.

(*d*) *Intrathoracic site.* An intrathoracic site is another possibility, and indeed I believe there is good evidence for the existence of a receptor which is capable of body water regulation. Sieker, Gauer and Henry (32, 33), and Drury, Henry and Goodman (14) investigated the effects of breathing against positive and negative pressure. They argued that the change in intrathoracic pressure resulting from these manœuvres would change the volume of blood in the low-pressure sites in the thorax. They found that breathing against a low pressure caused a water diuresis, and breathing against positive pressure an antidiuresis. These changes would be of the type expected in a regulator of body water. Thus, excess fluid retention would raise the volume of fluid in the pulmonary veins, and so lead to a diuresis. Diagrammatically, this system would regulate body water as follows (Figure 3):

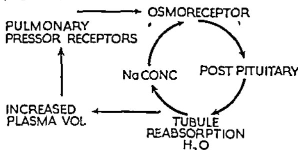


FIG. 3 Diagrammatic representation of a possible method of body water control.

However, this system cannot be the one we are seeking. Neither negative nor positive pressure-breathing affect sodium excretion, an invariable accompaniment of altered circulation.

Borst's Hypothesis

Borst has put forward an hypothesis as follows (4):

The compartment, whose volume is primarily controlled, is the vascular compartment.

Alterations in volume of fluid here cause alterations in filling pressure of the heart, and so changes in cardiac output. When the cardiac output is inadequate, water and salt are retained, so raising the filling pressure. You will note that this explains the occurrence of the reaction of circulatory inadequacy in cardiac failure.

This attractive hypothesis is very difficult to prove or disprove, for minor changes in cardiac filling pressure and cardiac output are difficult to demonstrate. In any event, it merely adds another link in a chain. How does the postulated change in cardiac output affect the receptor? It must be by some distant pressure or chemical change.

Thus, there is no convincing demonstration of the existence of a baroreceptor with the correct properties. Further search may reveal one.

Chemoreceptors

Changes in circulation could result from water and salt retention along the lines suggested by Borst, and just discussed. The changed circulation could then alter the rate of transport to, or removal from, the tissue of a substance whose tissue concentration would indirectly provide the signal of an altered state of hydration.

In the search for likely substances to affect a receptor we note the following. Cardiac failure of both the low and the high cardiac output type may be associated with fluid retention and the reaction of circulatory inadequacy (*12*). The factor common to both of these types of failure is an alteration in the transport of blood gases. Let us then investigate the possibility that altered blood-gas tension in tissues is the signal we are seeking.

Oxygen tension. If a change in tissue oxygen tension (pO_2) is the signal, we would expect that an 'inadequate' circulation would lower tissue pO_2, and that this would give rise to water and salt retention. Let us then examine the effect of lowering tissue pO_2 by breathing 10 per cent oxygen, or by placing the subject in a decompression chamber. Figure 4 demonstrates such a study.

There is a marked effect, but as may be seen, the very reverse of what we would expect of the regulating system.

CO_2 tension. In the same way, if pCO_2 in the tissues is the signal we would expect that a rise in pCO_2 would be associated with an inadequate circulation and fluid and salt retention. Let us then examine the effect of breathing 5 per cent CO_2. Again there is a marked effect (3). I will show it in some detail, as the technique of demonstrating the effect illustrates the type of control study that is needed in investigation of water and electrolyte problems.

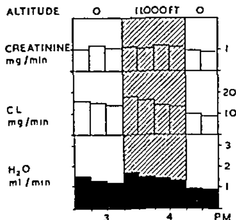


FIG. 4. The effect of lowered oxygen tension on excretion of creatinine, chloride and water. The subject was placed in a decompression chamber simulating an altitude of 11,000 feet (from Bull, unpublished).

Normal subjects, who were accustomed to self-experimentation, were investigated, because emotional disturbances may cause diuresis or antidiuresis. They were all studied in the afternoons so as to distinguish clearly a possible effect from the normal diurnal rhythm of water and salt excretion (5).

Dietary intake of fluid and minerals hardly varied from day to day during the study, and in particular, lunch was the same each day. One subject remained on an unvarying diet for two weeks during the course of the study. This is necessary, because changes in mineral and fluid content of the diet have profound effects on water and mineral excretion. Even the drinking of 1 litre of water 12 hours before may cause a change in pattern of excretion on application of a stimulus (21).

Alterations in volume of fluid here cause alterations in filling pressure of the heart, and so changes in cardiac output. When the cardiac output is inadequate, water and salt are retained, so raising the filling pressure. You will note that this explains the occurrence of the reaction of circulatory inadequacy in cardiac failure.

This attractive hypothesis is very difficult to prove or disprove, for minor changes in cardiac filling pressure and cardiac output are difficult to demonstrate. In any event, it merely adds another link in a chain. How does the postulated change in cardiac output affect the receptor? It must be by some distant pressure or chemical change.

Thus, there is no convincing demonstration of the existence of a baroreceptor with the correct properties. Further search may reveal one.

Chemoreceptors

Changes in circulation could result from water and salt retention along the lines suggested by Borst, and just discussed. The changed circulation could then alter the rate of transport to, or removal from, the tissue of a substance whose tissue concentration would indirectly provide the signal of an altered state of hydration.

In the search for likely substances to affect a receptor we note the following. Cardiac failure of both the low and the high cardiac output type may be associated with fluid retention and the reaction of circulatory inadequacy (12). The factor common to both of these types of failure is an alteration in the transport of blood gases. Let us then investigate the possibility that altered blood-gas tension in tissues is the signal we are seeking.

Oxygen tension. If a change in tissue oxygen tension (pO_2) is the signal, we would expect that an 'inadequate' circulation would lower tissue pO_2, and that this would give rise to water and salt retention. Let us then examine the effect of lowering tissue pO_2 by breathing 10 per cent oxygen, or by placing the subject in a decompression chamber. Figure 4 demonstrates such a study.

There is a marked effect, but as may be seen, the very reverse of what we would expect of the regulating system.

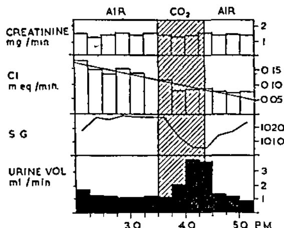


Fig. 6. The effect of breathing 5 per cent CO_2 on excretion of creatinine, chloride and water, and the urine specific gravity. Note that the gradual fall in chloride excretion, which was taking place during the afternoon, was unaffected by breathing CO_2 (from Barbour *et al.*, 1953).

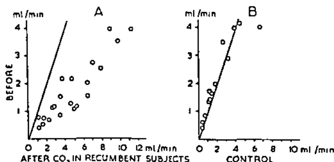


Fig 7 The changes in urine flow during control studies (B), and following the breathing of 5–7 per cent CO_2 (A) (from Barbour *et al.*, 1953).

The position of the subject studied was known and recorded. In most studies the subject was recumbent.

Control studies of the effect of laying down were made.

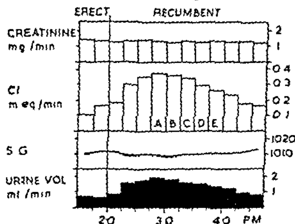


FIG. 5. *Control study* *The effect of lying down on the excretion of water, creatinine, chloride and urine specific gravity* (from Barbour *et al.*, 1953).

Experimental periods were always preceded by a further control period to ensure that urine flow was falling. The subjects then breathed 5 to 7 per cent CO_2 from large Douglas bags with low-resistance non-return valves. Figure 6 shows a typical result.

The over-all results on urine flow are shown in Figure 7.

The effects on excretion of urinary solutes were also studied.

TABLE 1. The Effect of Breathing 5% CO_2 on the Excretion of Various Substances
(Adapted from Barbour *et al*, 1953)

Substance	Change in excretion	*P*	Difference detectable at P = 0·05
Creatinine	− 1·4%	0·7 > P > 0·6	5·9%
Na	+ 4·0%	0·9 > P > 0·8	7·4%
Cl	− 1·1%	0·7 > P > 0·6	5·4%
K	+ 7·5%	0·1 > P > 0·05	8·6%
Total mols.	+ 34·2%	0·001 > P	14·8%
Urea	+ 18·9%	0·02 > P > 0·01	14·3%
Inulin clearance	− 14·2 ml/min.	0·2 > P > 0·1	20·8 ml/min
Effective Renal Plasma Flow	− 26·1 ml/min	0·5 > P > 0·4	71 ml/min.

However, on looking further into the matter, we find that hyperventilation causes a fall in CBF and a rise in urine flow and salt excretion (18, 37). See Figure 8.

Thus, we have found no easy answer in a study of the effects of altered blood-gas composition. Possibly some other indirect effect is important. There are still many other possibilities to investigate. Drs. Beck, Brennan and I are, at the moment, attempting to study the effects of alterations in oxidation-reduction potential on water and ion excretion. Eh would, in theory, be altered by altered CO_2 tension via a pH effect, and would also be altered by changes in pO_2. So far we have drawn a blank, but we are still feeling our way.

At all events you will see that a positive answer seems as far away as ever, but that there are still many lines of investigation to be followed.

SUMMARY

Following Peters and Borst, I have attempted to show that it is necessary to postulate the existence of a receptor system sensitive to changes in water content of the whole body, or part of it. Certain possible properties of this receptor are discussed.

REFERENCES

1 BADER, R. A., ELIOT, J. W and BASS, D. E. Renal and hormonal [illegible]

2. [illegible]

3. [illegible] dioxide on urine flow and mineral excretion. *Clin. Sci.* **12**, 1. (1953)

4. BORST, J G G. The maintenance of an adequate cardiac output by the regulation of the urinary excretion of water and sodium chloride; an essential factor in the genesis of oedema. *Acta med. Scand* **130**, Supp 207 (1948).

5 BORST, J G G and DE VRIES, L. A The three types of 'natural' diuresis *Lancet*, 1950, **ii**, 1

6 BRIGDEN, W., HOWARTH, S and SHARPEY-SCHAFER, E. P. Postural changes in the peripheral blood flow of normal subjects, with observations on vaso-vagal fainting reactions as a result of tilting, the

You will note that there is no significant change in excretion of Na and Cl. We are dealing with a pure water diuresis.

Thus, once more the direction of change in excretion is wrong for the postulated receptor. Furthermore, CO_2 does not affect mineral excretion.

Cerebral blood flow. These changes associated with altered pCO_2 and O_2 are real. Could it be that they are indirectly affecting a receptor in the following way?

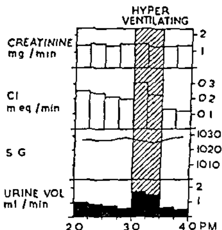


FIG. 8 The effect of hyperventilation on excretion of creatinine, chloride and water, and the urine specific gravity, in a normal recumbent subject (*from* Barbour *et al.*, 1953).

It so happens that the two most potent regulators of the cerebral blood flow (CBF) are the partial pressure of O_2 and CO_2 in arterial blood. For instance, Kety and Schmidt showed that breathing 5 per cent CO_2 caused approximately 75 per cent rise in CBF, and breathing 10 per cent O_2 a 25 per cent rise (19). Is it possible that something consequent on the rise in CBF is the signal we are seeking? When we examine the state of the cerebral circulation in various conditions, we find what at first sight is an encouraging correlation between fall in cerebral blood flow and water and salt retention, e.g. body-tilting (31), congestive cardiac failure (30), exercise and oligaemic shock (18) are all associated with a fall in cerebral blood flow, and water and salt retention.

However, on looking further into the matter, we find that hyperventilation causes a fall in CBF and a rise in urine flow and salt excretion (18, 37). See Figure 8.

Thus, we have found no easy answer in a study of the effects of altered blood-gas composition. Possibly some other indirect effect is important. There are still many other possibilities to investigate. Drs. Beck, Brennan and I are, at the moment, attempting to study the effects of alterations in oxidation-reduction potential on water and ion excretion. Eh would, in theory, be altered by altered CO_2 tension via a pH effect, and would also be altered by changes in pO_2. So far we have drawn a blank, but we are still feeling our way.

At all events you will see that a positive answer seems as far away as ever, but that there are still many lines of investigation to be followed.

Summary

Following Peters and Borst, I have attempted to show that it is necessary to postulate the existence of a receptor system sensitive to changes in water content of the whole body, or part of it. Certain possible properties of this receptor are discussed.

REFERENCES

1 Bader, R. A., Eliot, J W. and Bass, D E Renal and hormonal [illegible]

2. Ba [illegible]

3 Ba [illegible] d [illegible] n dioxide on urine flow and mineral excretion *Clin Sci* **12**, 1 [illegible] (1953).

4. Borst, J G G The maintenance of an adequate cardiac output by the regulation of the urinary excretion of water and sodium chloride, an essential factor in the genesis of oedema *Acta med Scand* **130**, Supp 207 (1948).

5. Borst, J G G. and de Vries, L. A. The three types of 'natural' diuresis *Lancet*, 1950, **ii**, 1

6 B[illegible] W Howarth S and S[illegible] [illegible]

lordotic posture, pregnancy and spinal anaesthesia. *Clin Sci.* **9**, 79. (1950).
7. BRUN, C., KNUDSEN, E. O. E. and RAASCHOU, F. The influence of posture on the kidney function. *Acta med. Scand.* **122**, 332 and 486. (1945).
8. BULL, G. M. Unpublished data.
9. CORT, J. Communication to Medical Research Society, 1953.
10. CULBERTSON, J. W., WILKINS, R. W., INGELFINGER, F. J. and BRADLEY, S. E. The effect of the upright posture upon hepatic blood flow in normotensive and hypertensive subjects. *J. clin. Invest* **30**, 305. (1951).
11. DARROW, D. C. and PRATT, E. L. Fluid therapy; relation to tissue composition and expenditure of water and electrolyte; Council on Foods and Nutrition. *J. Amer. med. Ass.* **143**, 365 and 432. (1950).
12. DAVIES, C. E. and KILPATRICK, J. A. Renal circulation in 'low output' and 'high output' heart failure. *Clin. Sci.* **10**, 53. (1951).
13. DEMPSTER, W J. and JOEKES, A. M. Functional studies of the kidney autotransplanted to the neck of dogs. *Acta med. Scand.* **147**, 99. (1953).
14. DRURY, D. R., HENRY, J. P. and GOODMAN, J. The effects of continuous pressure breathing on kidney function. *J. clin. Invest.* **26**, 945. (1947).
15. EPSTEIN, F. H., GOODYER, A. V. N., LAWRASAN, F. D. and RELMAN, A. S. Studies of the antidiuresis of quiet standing. *J. clin. Invest.* **30**, 63. (1951).
16 FISHMAN, R A The failure of intracranial pressure—Volume change to influence renal function. *J clin. Invest.* **32**, 847. (1953).
17. FITZHUGH, F W Jr, MCWHORTER, R L. Jr., ESTES, E. H. Jr., WARREN, J. V. and MERRILL, A. J. The effect of application of tourniquets to the legs on cardiac output and renal function in normal human subjects *J. clin Invest* **32**, 1163. (1953).
18. KETY, S. S. Circulation and metabolism of human brain in health and disease *Amer. J Med* **8**, 205. (1950).
19. KETY, S. S and SCHMIDT, C. F. The effects of altered arterial tensions of carbon dioxide and oxygen on cerebral blood flow and cerebral oxygen consumption of normal young men *J. clin. Invest* **27**, 484. (1948)
20. KLISIECKI, A, PICKFORD, M, ROTHCHILD, P and VERNEY, E B. Absorption and excretion of water by mammal; relation between absorption of water and its excretion by innervated and denervated kidney. *Proc. roy Soc. B*, **112**, 496 (1933).
21. LADD, M. Effect of prehydration on response to saline infusion in man. *J. appl Physiol.* **3**, 379 (1951).
22. LEWIS, J M Jr., BUIE, R M., SEVIER, S M and HARRISON, T. R. Effect of posture and of congestion of head on sodium excretion in normal subjects. *Circulation*, **2**, 822 (1950).

23 MAXWELL, M. H., BREED, E. S. and SCHWARTZ, I. L. Renal venous pressure in chronic congestive heart failure. *J. clin. Invest.* **29**, 342. (1950).

24. MOKOTOFF, R. and ROSS, G. Effect of spinal anaesthesia on renal ischaemia in congestive heart failure. *J. clin. Invest.* **27**, 335. (1948).

25. MYERS, J. D. and HICKAM, J. B. An estimation of the hepatic blood flow and splanchnic oxygen consumption in heart failure. *J. clin. Invest* **27**, 620 (1948).

26 McMICHAEL, J. and SHARPEY-SCHAFER, E. P. Cardiac output in man by direct Fick method; effects of posture, venous pressure change, atropine and adrenaline. *Brit Ht J.* **6**, 33 (1944).

27. PETERS, J. P. Role of sodium in production of oedema *New Engl. J. med.* **239**, 353 (1948).

28. ROBINSON, J R. *Osmo regulation in surviving slices from the kidneys of* adult rats *Proc roy Soc. B*, **137**, 378. (1950)

29. ROSENBAUM, J. D, PAPPER, S and ASHLEY, M M. Variations in renal *tubular reabsorption of sodium, independent of change in adreno-*cortical hormone. *J. clin Invest* **31**, 657 (1952)

30. SCHEINBERG, P. Cerebral circulation in heart failure. *Amer. J Med.* **8**, 148. (1950).

31 SCHEINBERG, P and STEAD, E. A Jr. The cerebral blood flow in male subjects as measured by the nitrous oxide technique Normal values for blood flow, oxygen utilization, glucose utilization, and peripheral resistance, with observations on the effect of tilting and anxiety. *J. clin Invest.* **28**, 1163 (1949)

32. SIEKER, H O, GAUER, O. H and HENRY, J P The effect of negative pressure breathing on renal function *J clin Invest* **31**, 662. (1952).

33 SIEKER, H. O, GAUER, O. H and HENRY, J. P. The *effect of con*tinuous negative pressure breathing on water and electrolyte excretion by the human kidney *J clin Invest* **33**, 572 (1954).

34. SMITH, H. W Chapter 22, Review, p. 663, 'Structure and function in health and disease, in *The Kidney*, Oxford Univ Press, New York. (1951).

35 SMYTHE, C McC, NICKEL, J F and BRADLEY, S. E The effect of epinephrine (USP), l-epinephrine and l-norepinephrine on *glomerular filtration rate, renal plasma flow and the urinary ex*cretion of sodium, potassium and water in normal man *J clin Invest* **31**, 499 (1952)

36. SQUIRE, J R Nephrotic Syndrome *Brit med. J* 1953, **II**, 1389.

37 STANBURY, S. W and THOMSON, A. E The renal response to respiratory *alkalosis Clin Sci* **11**, 357 (1952)

38 STARLING, E H Physiological factors involved in the causation of dropsy *Lancet*, 1896, **i**, 1267, 1331, 1407

39 VERNEY, E B Agents determining and *influencing the functions of* the pars nervosa of the pituitary *Brit. med J* 1948, **II**, 119.

40. DE WARDENER, H. E. and McSWINEY, R. R. Renal haemodynamics in vasovagal fainting due to haemorrhage *Clin. Sci.* **10**, 209. (1951).

41. WARREN, J. V., BRANNON, E. S., STEAD, E. A Jr. and MERRILL, A. J Effect of venesection and pooling of blood in extremities on atrial pressure and cardiac output in normal subjects, with observations on acute circulatory collapse in three instances. *J. clin. Invest* **24**, 337. (1945).

42. WARREN, J. V. and Stead, E. A. Jr. Fluid dynamics in chronic congestive heart failure. *Arch. int. Med.* **73**, 138. (1944).

43 WELT, L. G., SELDIN, D. W., NELSON, W. P., GERMAN, W. J and [illegible] metabolism and [illegible]

44. W[illegible] nd general effects of large, rapid plasma infusions in convalescent men. *J. clin. Invest.* **29**, 251. (1950).

XIV

Reactions to Bacterial Invasion

A. A. MILES

The invasion of a host by a bacterium, whether it results in acute or chronic disease, is a process which may be said to start when the bacterium comes into contact with some external tissue, and to finish with the death of the host; and the reactions to this invasion could cover every deviation from the healthy norm that the clinician can recognize or the pathologist devise a test for. Invasion proper starts with the establishment of the bacteria within the tissues, when a breach, either mechanical or functional, has been made in the skin or one of the mucosal surfaces. It is highly probable that such breaches in the superficial defences occur many times a day in the life of an animal; and the events leading to it we may distinguish as the process of infection. Successful infection depends on the continued survival of the potentially pathogenic bacteria that reach a particular site on the epithelial surface of the host, until a favourable accident happens at that site.

In this early infective stage the tissue is reacting only in the negative sense of submitting to destruction. During the invasion also the tissues submit to destruction; but I propose to confine this discussion to reactions in which there is active, and generally speaking defensive, adaptation to the invader. At this point we may usefully borrow a term from the microbial enzymologists, and contrast these adaptive defences with constitutive defences. The constitutive enzyme, it will be recalled, is produced by a cell no matter what food is supplied, whereas the adaptive enzyme is produced only in response to certain stimuli. By analogy, the constitutive defences, such as the antibacterial

lysozyme of the nasal mucosa, are in full working order independently of the presence of an invader. The inflammation round a local invasion of bacteria, on the other hand, is clearly adaptive. As we might expect in the reactions of metazoan animals, the adaptations, unlike the enzyme adaptations of a single cell, may persist for some time after withdrawal of the stimulus. We shall find too, as with enzymes, that some constitutive defences, normally maintained at a relatively low level, are capable of adaptive enhancement.

After nearly a century of pathology and bacteriology, a wide variety of responses to invasion, both cellular and humoral, has been firmly established. But, their interrelations, their relative values, and their susceptibility to improvement in the infected animal, are much less clear; and it is these features I propose to examine.

[illegible] to Invasion

[illegible] sponse is, of course, the formation of specific [illegible] ntibody response is doubly adaptive. In the first place the newly arriving bacterial antigen stimulates existent antibody-forming cells, and in the second place, it stimulates the multiplication of these cells. It has become abundantly clear in the last twenty years that this cellular response occurs not only in places like the spleen and lymphatic tissue, but also in antibody-forming cells that have proliferated locally. The now classic experiment of McMaster and Hudack (1935) proved that lymph nodes of mice draining the site of inoculation of antigen could produce antibody more abundantly than the rest of the body. Thus, salmonella bacilli were injected into one ear, antigenically unrelated *Chromobacterium prodigiosum* into the other; and after seven days the regional lymph nodes on the salmonella side contained more salmonella agglutinins than the contra-lateral nodes, whereas chromobacterium agglutinins predominated on the chromobacterium side. The object of the double injection was to produce inflammation in the nodes of both sides, so as to exclude the interpretation of a high antibody content in either of them as due merely to inflammatory exudation of serum antibodies. It is now evident that antibody is also formed by cells adapta-

tively proliferating in the connective tissues at the site of injection of antigen (see Oakley *et al.*, 1950). Lymphocytes, plasma cells or their immediate precursors, macrophages, and combinations of these cells have been postulated as the source of this antibody. The balance of evidence is perhaps in favour of the plasma cell, but for our purpose it is sufficient to note that all these cells belong to the class of connective tissue cells described in 1937 by Taliaferro and Mulligan as the lymphoid-macrophage system; a system that includes not only the reticulo-endothelial system of Aschoff, but all Metchnikoff's macrophages as well.

The value of the antibody response is well established. It intervenes within a few days of infection to sensitize the invader to phagocytosis or to lysis by humoral factors, and perhaps to immobilize it; to neutralize some of its toxins; and in some cases to interfere directly with its growth, apparently by inhibiting oxidative processes. Moreover, whether as the result of manifest disease, of sub-clinical infection, or of deliberate vaccination, the adaptation to the antigens of the invader is in many cases deeply enough impressed on the host to provide a durable specific immunity. For a time at least the acquired immunity is constitutive (though not *genetically* so) in the sense that persisting antibody is immediately available to curb invasion. More important still, the antibody-forming cells have acquired an enormously increased capacity to respond rapidly to a fresh experience of the same antigen.

The triumphs of prophylactic immunization and the lesser, but definite triumphs of antiserum therapy, attest the value of the antibody responses. But as a defence mechanism, it has two striking limitations. Antibodies are effective mainly outside the body cells; once inside a cell a bacterium appears to be unaffected by them. The same applies to habitually intracellular parasites like the viruses; although the extracellular antiviral antibody may be useful, because it can combine with an invading virus that is on its way to the susceptible cells. The other limitation is the allergy that may accompany the antibody response. It would be overbold for anyone to be dogmatic about its benefits or disadvantages in long-lasting diseases, like tuber-

culosis, in which 'bacterial' allergy usually appears. In rheumatic fever there is a rather better case for believing that if we could suppress the allergic response to *Strep. pyogenes* or to products of the host's tissues modified by *Strep. pyogenes*, and leave an uncomplicated immune response to aid in the elimination of the invader, we should benefit that substantial proportion of the human race which suffers from rheumatic fever. The immune response, in fact, does not appear to be consistently directed to the end of preserving the body from all ills; and a diehard teleologist who wished to maintain that it is so directed ... back upon the excuse—so damaging to the ... fails to be ... it means well.

As an adaptive mechanism the antibody-forming apparatus is enormously tough. It is usually only after damage of the degree produced by starvation, deprivation of vitamins to the point where the animal does not want to feed, or by doses of X-rays that smash the lymphoid-macrophage system, that the antibody response can be significantly impaired (Taliaferro and Taliaferro, 1951). Its much discussed, non-specific stimulation by hormones is still in doubt.

Non-specific Adaptation to Invasion

Let us now turn to the apparently non-specific response. In recovered or recovering hosts, the invader is killed and either ... and lysis, and perhaps diges- ... ut the close association of both intracellular and ... death with the granular leucocytes—the microphages of Metchnikoff—and with the free and fixed cells of the lymphoid-macrophage system, justifies our starting with cellular adaptations. This mesenchymatous cellular system is obviously constitutive, and its constitutive nature is nowhere clearer than in its behaviour towards intravenously injected dyes and inert particles, or towards bacteria. The uptake of such particles from the circulation begins immediately and is for several hours a continuing function of the fixed macrophages of the reticulo-endothelial

system, which are abundant in the spleen, liver, lymph nodes, and bone marrow. The system is also adaptive. For example, in many attempts at reticulo-endothelial blockade by massive doses of particles intended to stuff the cells to repletion, the result has been a stimulation, not a depression, of the clearing powers. This adaptive effect and its manifestation within an hour or so was neatly demonstrated by Halpern and his colleagues (Biozzi *et al.*, 1953), in a recent study of the kinetics of clearance of carbon particles. The rate of clearance proved to be determined by the ratio of the amount injected to the amount already in the cells; and it decreased as the cells became saturated. Nevertheless, it fell less rapidly with large than with small doses; in other words, the bigger dose had stimulated the greater cellular activity.

Living bacteria are at first taken up in the same way as inert particles and an experimental bacteraemia disappears rapidly. But the subsequent course of infection depends on the power of the cells to suppress growth of the bacteria they have ingested, or to resist the bacterial toxins. When it is high, as in a naturally resistant or an immune animal, digestion occurs; when it is low, a further, and fatal, bacteraemia may supervene.

The leucocytes, complement and other antibacterial factors of the normal blood are also constitutive, but their passage into inflamed tissues is adaptive; and so is the production of leucocytes in the haemopoietic organs stimulated by the bacterial invasion. A more dramatic adaptation is the rapid transformation, in a locally inflamed region, of fixed tissue cells and of lymphocytes and monocytes from the blood stream, into phagocytic and often visibly free macrophages. It is enough to record that this is a well-established phenomenon, and to leave the still-vexed question of the pedigree of the various cells to the haematologist (see Rebuck, 1947).

Like the capacity for antibody response, the clearing mechanism, and the resistance to fatal infections depending on its integrity, are destroyed or significantly diminished only when the cellular system is severely damaged—by massive, near-lethal doses of X-rays, by reticulo-endothelial blockade, or by what is best described, in view of the quite unphysiological

culosis, in which 'bacterial' allergy usually appears. In rheumatic fever there is a rather better case for believing that if we could suppress the allergic response to *Strep. pyogenes* or to products of the host's tissues modified by *Strep. pyogenes*, and leave an uncomplicated immune response to aid in the elimination of the invader, we should benefit that substantial proportion of the human race which suffers from rheumatic fever. The immune response, in fact, does not appear to be consistently directed to the end of preserving the body from all ills; and a diehard teleologist who wished to maintain that it is so directed would have to fall back upon the excuse—so damaging to the excused in ordinary life—that if the immune response fails to be obviously beneficial in all its allergic manifestations, at least it means well.

As an adaptive mechanism the antibody-forming apparatus is enormously tough. It is usually only after damage of the degree produced by starvation, deprivation of vitamins to the point where the animal does not want to feed, or by doses of X-rays that smash the lymphoid-macrophage system, that the antibody response can be significantly impaired (Taliaferro and Taliaferro, 1951). Its much discussed, non-specific stimulation by hormones is still in doubt.

Non-specific Adaptation to Invasion

Let us now turn to the apparently non-specific response. In recovered or recovering hosts, the invader is killed and either mummified or digested. Its death and lysis, and perhaps digestion, may be extra- or intra-cellular. But the close association of both intracellular and extracellular death with the granular leucocytes—the microphages of Metchnikoff—and with the free and fixed cells of the lymphoid-macrophage system, justifies our starting with cellular adaptations This mesenchymatous cellular system is obviously constitutive, and its constitutive nature is nowhere clearer than in its behaviour towards intravenously injected dyes and inert particles, or towards bacteria. The uptake of such particles from the circulation begins immediately and is for several hours a continuing function of the fixed macrophages of the reticulo-endothelial

system, which are abundant in the spleen, liver, lymph nodes, and bone marrow. The system is also adaptive. For example, in many attempts at reticulo-endothelial blockade by massive doses of particles intended to stuff the cells to repletion, the result has been a stimulation, not a depression, of the clearing powers. This adaptive effect and its manifestation within an hour or so was neatly demonstrated by Halpern and his colleagues (Biozzi *et al.*, 1953), in a recent study of the kinetics of clearance of carbon particles. The rate of clearance proved to be determined by the ratio of the amount injected to the amount already in the cells; and it decreased as the cells became saturated. Nevertheless, it fell less rapidly with large than with small doses; in other words, the bigger dose had stimulated the greater cellular activity.

Living bacteria are at first taken up in the same way as inert particles and an experimental bacteraemia disappears rapidly. But the subsequent course of infection depends on the power of the cells to suppress growth of the bacteria they have ingested, or to resist the bacterial toxins. When it is high, as in a naturally resistant or an immune animal, digestion occurs; when it is low, a further, and fatal, bacteraemia may supervene.

The leucocytes, complement and other antibacterial factors of the normal blood are also constitutive, but their passage into inflamed tissues is adaptive; and so is the production of leucocytes in the haemopoietic organs stimulated by the bacterial invasion. A more dramatic adaptation is the rapid transformation, in a locally inflamed region, of fixed tissue cells and of lymphocytes and monocytes from the blood stream, into phagocytic and often visibly free macrophages. It is enough to record that this is a well-established phenomenon, and to leave the still-vexed question of the pedigree of the various cells to the haematologist (see Rebuck, 1947).

Like the capacity for antibody response, the clearing mechanism, and the resistance to fatal infections depending on its integrity, are destroyed or significantly diminished only when the cellular system is severely damaged—by massive, near-lethal doses of X-rays, by reticulo-endothelial blockade, or by what is best described, in view of the quite unphysiological

doses employed, as chronic cortisone poisoning (see Kass and Finland, 1953). They are diminished under exposure to lesser experimental rigours like continued acidosis (Egoroff *et al.*, 1935) or continued ethanol intoxication (Klepser and Nungester, 1939). The system is, in fact, as tough and resilient as the antibody-forming apparatus.

Although there are contradictory details in parts of it, the picture of the cellular defences is one of mesenchymal reserve tissues, mainly lymphoid, that adaptively give rise to the blood cells, the local phagocytes and the connective tissue repair cells (Taliaferro, 1949). The sequence in which these cells appear in the inflammatory reaction, and the histological picture of an invasion at any stage, varies with the kind of invader. It has indeed no necessary relation to special defence mechanisms required in a given disease. In all probability the chief determining factors are the species and virulence of the infecting bacterium; each species has its own peculiar cytotoxins with predilections for certain tissue cells; and the virulence determines the acuity of the attack and hence of the response.

The Role of Antibody in 'Non-specific' Defence. We now have in broad perspective the two great defensive systems. First, the antibody-forming apparatus, which is capable of an entirely new and specific adaptive response; and second, the cellular system which is constitutive, but capable of ready adaptive improvement as an ingester and a digester of the invader—an adaptation that is apparently non-specific. The broad relation of the two is easily defined. The defensive cellular reactions—immobilization, killing, ingestion, removal and digestion of the invader—in the specifically immune animal differ only quantitatively from those in an animal without detectable antibody; they take place more rapidly and more efficiently. Francis (1950) made the same point when he recently wrote: 'It is a constantly impressive phenomenon that the normal mechanisms seek to accomplish the same ends that are specifically completed by the immune process.' But you will note he speaks of a *completion* by the immune response. This is obviously true late in the course of an infection. But when we come to evaluate the

different reactions, we must ask whether specific antibody may not *always* be necessary for their completion. In other words, are all successful defence responses by the phagocytic component of the cellular systems mediated by antibody produced at the same time?

At first sight, remembering that antibodies are seldom detectable in the blood of a newly inoculated animal until after the lapse of a few days, we might safely label the reactions in the first few days of invasion as entirely non-specific. But we must take into account local defence; and the possibility of local formation of antibody in quantities far too small to appear in the circulation. Because, since phagocytic macrophages may arise from tissue cells within an hour of invasion (Tompkins and Grillo, 1953), it is possible that antibody-forming cells do so. We might argue that such cells, even if quickly formed, would take time to respond to antigen. What we know of antigens points to their being macromolecular substances that do not slip easily into the metabolic machine, either for digestion or other use. There is thus no good reason why a cell should not begin to form antibodies as soon as it has failed to digest the invader beyond the macromolecular stage; and visible digestion of intracellular bacteria has been observed within a few hours of their ingestion. With the antigenic stimulus provided by a lethal dose of virulent pneumococci in the peritoneal cavity, N [illegible]

3–4 hours of stimulation. In any analysis, therefore, of the comparative value of the non-specific cellular and humoral reactions we must bear in mind the possibility of a specific element after the first few hours.

Constitutive Cellular Defences

Starting our analysis with the constitutive cellular defences, we have some evidence from experimental genetics which, at first sight, offers a means of evaluating them. By inbreeding mice, hens or rabbits, selected for their resistance or susceptibility to a given infection, strains of the animals have been produced in

which the results of back-crosses strongly suggest that the resistance is genetically determined. When resistant strains are compared with susceptible, they are found to differ significantly in a number of features. They may have, for example, a greater number of circulating leucocytes; a slightly higher body temperature and a greater capacity to resist chilling; a higher pH of the blood; a greater capacity of the reticulo-endothelial cells of the spleen and liver to digest the invader, and of the liver parenchyma to resist bacterial intoxication (see Gowen, 1948). Table 1 summarizes the characters of two mouse strains,

TABLE 1. Inbred Mouse Strains selected for Resistance to a Salmonella *Infection* (from Weir, 1949)

	Strain	
	S	L
Mean blood pH	7·427	7·355
Mean WBC per cu. mm.	19,600	11,600
Per cent Ly[1]	73	75
Per cent Gr[1]	17	15
Per cent surviving 200,000 oral *S. typhi-murium*	86	13

[1] Ly = lymphocytes, Gr = granulocytes.

S and L, of high and low resistance to *Salmonella typhi-murium*, and the associated white blood counts and blood pH (Weir, 1949). It is gratifying to note that the resistant strains differ in just those features we should associate with the efficient killing of bacteria But this genetic evidence unfortunately does not take us very far. When the mice are selected, not for resistance to the Salmonella disease, but for those features which appear to be associated with resistance, we get very different results. The mouse strain T, in Table 2, of a given *susceptibility to Salmonella typhi-murium*, was selected for high and low leucocyte count (LC/H and LC/L), and high and low blood pH (pH/H and pH/L). Subsequent resistance tests revealed a directly opposite effect; the more resistant animals had the lower count and the lower blood pH (Weir *et al.*, 1953). In other words, in the S and L strains the particular characters were associated

with resistance as a result of one set of interactions with unknown, and presumably genetically determined, factors. The different mode of selection in another strain of mice has obviously produced different interactions. The association of,

TABLE 2 Inbred Mouse Strains selected for pH and Leucocyte Content of Blood (from Weir *et al.*, 1953)

	Strain[1]				
	T	LC/H	LC/L	pH/H	pH/L
Mean blood pH	7·45			7·47	7·42
Mean WBC per cu mm.	8,700	15,500	6,700		
Per cent surviving 200,000 oral *S. typhi-murium*	43	16	27	28	43

[1] For explanation of symbols, see text, p 242

say, a high leucocyte count with resistance in the S strain cannot, therefore, be used as independent evidence of the value of a high leucocyte count. The association is only *consistent* with results of other experimental studies. A great deal more genetic analysis will be needed to produce really independent evidence of the value of a particular defence mechanism.

Adaptive Cellular Defences

It will be convenient to consider the adaptive reactions in three stages of the invader's progress from epithelium to the generalized disease. The first, when infection has taken place, and the bacterium is in the tissues, may be called *primary lodgement* and this may, of course, remain local, though progressive. On the other hand, it may proceed *via* the lymphatic channels to the regional lymph nodes, where we have the second possible hold-up, the *lymphatic lodgement*. The third stage is the *systemic*, with access to the phagocytic reticulo-endothelial cells, and dissemination among various organs.

Defence at the Primary Lodgement. Experimentally, the defences against the primary lodgement may be modified either locally

or generally. Local resistance is destroyed when the haematopoietic system is poisoned by benzol; or, more informatively, when only the circulating granulocytes are removed by treating an animal with anti-granulocyte serum. Most of these reagents, however, act slowly and too ponderously to give the experimental freedom necessary to determine the time-course of the early reactions to invasion; and we can get more information out of local modifications.

As early as 1898, Cobbett and Melsome demonstrated the increased resistance of the rabbit's ear to streptococci 24 hours after the injection of mustard oil. But experiments of this kind tell us only of the efficacy of responses that are 24 hours old. Modifying agents like mustard oil, whose action is difficult to control, are much too crude for exploring the adaptive resistance in this way. We need more delicate tools. One such tool we have in the vasoconstrictor adrenaline. On intracutaneous injection, it persists long enough in the tissue to blanch the skin for 15 minutes to 6 hours according to the dose; and after recovery from moderate doses of 2–4 μg. producing a 2–3 hour ischaemia, the skin shows no signs of abnormality, either histologically or in its reactions to infection. In these doses, moreover, adrenaline has no action on chemotaxis of leucocytes, or phagocytosis, or on the bactericidal power of the blood; and it has no direct action on the bacteria. It appears, therefore, to act solely by cutting the blood supply.

Given directly with the bacteria, 2 μg. enhance the local infection to a degree that depends on the species and strain of pathogen used. When 10-fold dilutions of a suspension of washed living bacteria are titrated in the skin, the enhancement can be measured by matching the number of bacteria required to produce the same-sized lesion with and without the drug. Figure 1 shows the enhancement of *Ps. pyocyanea.* The shaded columns represent the diameter of the central necrotic area and the white columns that of the inflammatory lesion measured at 24 hours. The adrenaline makes 10,000 bacilli behave like a million injected without it. On this basis, the reaction to the primary lodgement is worth a 99 per cent kill to the invaded host. Enhancement measured in this way proves to be from 10-

to 100-fold with group C streptococci and staphylococci, 100- to 1,000-fold with *Escherichia coli* and *Ps. pyocyanea* and the diphtheria bacillus, and 100,000- to a million-fold for *Cl. welchii* (Evans *et al.*, 1948). In other words, the early reaction in the primary lodgement may have a very high protective value.

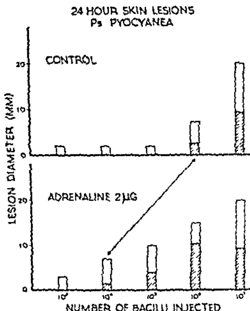


FIG. 1. The 100-fold enhancement of washed living *Ps. pyocyanea* in the skin of guinea-pigs by 2 μg. adrenaline in an injection volume of 0·1 ml The white column indicates the mean diameter of the inflammatory areas, the shaded column that of the central necrotic areas.

A slight modification of this intracutaneous titration enables us to estimate the time relations of the bactericidal response. Bacteria are injected first, the hole made by the needle being carefully marked with dye, and the adrenaline is later superinjected into the same site. Figures 2 and 3 show two results with a group C streptococcus in guinea-pigs and with vaccinia virus in rabbits. In both, the adrenaline effect declines, and after 3–4 hours it fails to enhance. This is characteristic also of infections with *Staph. aureus*, *Ps. pyocyanea*, coliform organisms, *Cl. welchii*, and the diphtheria bacillus. Most of the lesions were

measured at 24 hours, and the vaccinia lesions, it should be noted, were read at 4 days. It follows that the reactions inhibited by adrenaline are complete within four hours, and, as far as the local lesion is concerned, within that period they decide the maximum development that will be observed from 1–4 days later. These decisive reactions affect even remoter consequences, such as death from generalized infections. For example, in some

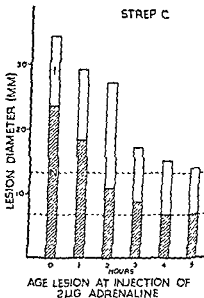


Fig.
strept
ages.
the shaded column that of the necrotic areas after 24 hours; the dotted lines indicate the respective mean diameters of untreated lesions.

experiments of Dr. A. C. Dutton (unpublished), the mortality of mice after 5 days or more, from a subcutaneous Strep. C *infection, is enhanced over 10-fold by local adrenaline* (Figure 4). But adrenaline superinjected into the local lesion at 4 hours scarcely affected the later mortality.

These 4-hour lesions have not become insensitive to adrenaline because of the inflammatory changes, since the blood supply of infectively inflamed tissues is visibly diminished by the drug. We must conclude that a decisive reaction to the primary

lodgement, with far-reaching consequences if the infection later generalizes, occurs during this critical period of four hours. The same decisive period is indicated when we diminish the blood supply by other means, such as the general shock (Miles and Niven, 1950) produced after two hours by intravenous bacterial endotoxin. When, and only when, shock is sufficient to bring the pressure and flow of blood in the peripheral small arteries of the

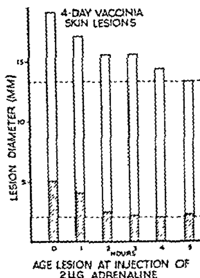
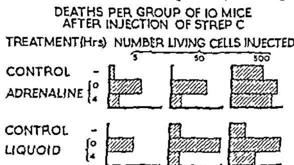


FIG. 3. The enhancement in rabbits of a constant dose of vaccinia virus by 2 μg adrenaline superinjected into skin lesions of various ages. Mean lesion-diameters indicated as in Figure 2, except that they were measured on the fourth day.

skin below a certain level, is there a striking enhancement of the lesion, as measured after 24 hours. But in the sites where the infection was 2–3 hours old by the time the shock had developed, the lesions were no bigger than those in control animals. This effect is not peculiar to generalized bacterial intoxication, because it occurred with shocking agents such as intraperitoneal hypertonic glucose solution and adenosine triphosphate. Now with both local adrenaline and general shock there appears to be no change in the permeability of the vessels; but a deficiency in the supply of circulating leucocytes, and of intravascular

pressure to push them through the vessel wall into the region of the primary lodgement. Microscopically, there is in 2-hour-old lesions enhanced by shock or adrenaline, a complete absence of tissue leucocytosis.

Deprivation of serum factors must also play a part. Our evidence for this is more indirect. The part played by one important humoral factor, complement, can be estimated by using the anti-complementary substance 'Liquoid' (sodium polyanethol



[illegible] graded doses of a group

sulphonate) in the place of adrenaline (Figure 5). Like adrenaline, liquoid enhances skin lesions, and here again the 4-hour period is decisive. But the figure shows two differences; adrenaline and liquoid do not affect the different bacteria to the same degree, suggesting that they interfere with different factors in defence; and though at 3–4 hours the lesions are insusceptible to liquoid, as they grow older they become susceptible again. The explanation of this late effect is not clear but it is possible that in 5–6 hours there is local antibody formation, or, as Maaloe (1948) suggested, there is a local formation of complement. Since normal phagocytosis and the bactericidal and opsonic effects of antibody are both mediated by complement, either explanation would be consistent with the late *liquoid* effect. You will note that the decisive period determined by liquoid in mice with a generalizing streptococcal infection, is also 4 hours (Figure 4).

The definition of this 4-hour decisive period in the primary lodgement enables us to assess the value of another reaction, whose defensive value had been advocated for many years by Menkin (1940, 1948); namely the prevention of systemic spread from a lesion, by thrombosis of the lymphatic channels. It is

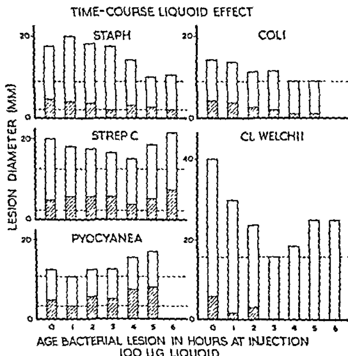


FIG. 5 The enhancement in guinea-pigs of a constant dose of five bacteria —*Staph. aureus*, a group C streptococcus, *Ps pyocyanea*, *E. coli* and *Cl. welchii*—by 100 µg. liquoid superinjected into skin-lesions of various ages. Mean lesion-diameters after 24 hours indicated as in Figure 2.

easy to show with necrotizing doses of severe inflammatory [illegible]

injection of ink into the lymphatic plexus of the guinea-pig's ear. Plate V, Figure 6a[1] shows a segment of the lymphatic plexus

[1] This plate will be found at the end of the book

in the ear of a normal guinea-pig. When the ink is injected into an ear with a day-old subcutaneous micro-abscess due either to *Staph. aureus* or *Strep. pyogenes*, a few millimetres in diameter, the plexus is sometimes obviously blocked at the centre of the abscess (*Plate V*, Figure 6*c*) where there is evident coagulative necrosis. Sometimes the abscess cavity is filled with ink (Plate V, Figure 6*b*) suggesting either little or no thrombosis in the lymphatic channels involved at the edge of the lesion. It can be shown that the ultimate size of these abscesses, like those in the skin of the trunk, is decisively determined by the reactions in the first four hours. In this period, therefore, the local spread of the infection so as to involve more and more lymph channels, with a consequent increase in the risk of a generalized disease, is to a large extent decided; and any *effective* lymphatic blockade should also be observable. In a series of tests on 24 streptococcal and 24 staphylococcal abscesses four hours old, we have never observed any blockage. The channels, even the smallest, were as patent as those in Figure 6*a*. At least in the earlier stages of the primary lodgement, therefore, there is no good reason to suppose that lymphatic blockade is either a usual or an effective defence reaction.

Next we must enquire whether in the primary lodgement, or indeed in any local reaction to invasion, the inflammatory response includes the manufacture of any bactericidal substances besides antibody and complement. Within recent years, a number of substances isolable from animal tissue have been described (Table 3). They include tissue polypeptides, sometimes highly basic, active against staphylococci, streptococci, spore-bearing bacilli like *B. anthracis*, and the tubercle bacillus. Basic proteins such as histones, protamines, and spermine activated by spermine oxidases, and fatty acids, are active against tubercle bacilli. Some of the substances are found only in certain organs—like thyroid, thymus or kidney, a distribution that may possibly explain the natural resistance of these organs to damage during a generalized infection by a particular bacterium. Some are found only in certain species of animal, and here again their presence may one day be correlated with species resistance. Their activity is demonstrable *in vitro* and is

mainly bacteriostatic. Sanguenin, the polypeptide from red blood cells, is the only one reported to be curative *in vivo*; and it is perhaps significant that this was isolated after relatively brutal hydrolysis of the cells (Micks *et al.*, 1951). But, as most of those who have described these substances are careful to insist, their participation in natural defences remains to be proved.

TABLE 3. The Antibacterial (mainly Bacteriostatic) Activity of some Mammalian Tissue Components

Source	Nature	Bacteria affected	Ref.[1]
Serum (human)	Globulins	*Br. abortus*	1
Brain, Spleen	Extract	*Staph. aureus*	2
Caecum, Pancreas	Basic Polypeptide	*B. anthracis*	3
Thymus (calf)	Histone		
Hydrolysed R.B C.	Polypeptide	*Strep. pyogenes*	4
Thyroid (pig)	Polypeptide	*E. coli* *Strep. pyogenes*	5
Kidney (guinea-pig)	Spermine, Spermine Oxidase	*Myco. Tuberculosis*	6
? Mammalian Cell Nuclei	Protamine and Histones		6
? Tissues	Aliphatic Acids		7
? Leucocytes	Lysozyme		8
Spleen, Liver	Washings or Extracts		9
Urine	?		10, 11

[1] References 1 Braun (1949), 2. Nutini and Lynch (1946), 3. Bloom *et al.* (1947), 4 Micks *et al* (1951), 5 Bloom and Prigmore (1952); 6 Hirsh (1953); 7 Dubos (1950), 8 Myrvik and Weiser (1951); 9 Bloom *et al* (1953), 10 Bjornesjö (1952), 11 Soltys (1953)

This approach is one facet of the biochemical investigation of the inflammatory reaction itself. Menkin (1940, 1948, 1949), for example, has described a number of factors in inflammatory exudates which he named according to their biological activity. They include leukotaxine, a polypeptide resulting from the breakdown of proteins, which increases capillary permeability, and in the tissue is chemotactic for leucocytes; a cytotoxic factor named necrosin; pyrexin, which induces fever; and two leucocytosis-promoting factors, apparently α-globulins, which stimulate an increase in circulating leucocytes. Some recent work of our own (Mackay *et al.*, 1953) suggests that it may not be necessary to provoke the biochemical consequences of acute and massive inflammation to obtain substances of similar bio-

logical activity. There are in the plasma or serum of the guinea-pig, in an inactive form, globulins, which when activated, are vasodilators, increase capillary permeability, change the vascular endothelium so that circulating leucocytes stick to it, and promote tissue leucocytosis. These globulins, though they are best prepared by a combination of ethanol and electrophoretic fractionation of the serum, when they are isolated in active form, can be activated by simple dilution of the serum. The easy activation of these blood-borne substances suggests that they may partake in the immediate reactions to the microbial or other traumata that act as trigger mechanisms in defensive inflammation. From their susceptibility to soya bean trypsin inhibitor, they appear to be proteases, though they are not fibrinolysins. Ungar (1950) has recently postulated activation of serum fibrinolysins as the essence of the inflammatory reaction, with substances like histamine, Lewis's H-substance, and leukotaxine as the active end-products. But here again we are in no position to assign a definitive role to any of these substances or factors as *necessary* chemical mediators in local inflammatory defence.

I have spent some time on the primary lodgement, and properly so, because it is the first line of defence against invasion. I have only brief comments to make on the next two stages, the lymphatic and the systemic.

Defence by lymph nodes. The essential facts of the lymph node reactions, are first that bacteria readily pass directly to the blood stream through non-inflamed nodes; secondly, bacteria are held up in inflamed lymph nodes, either in phagocytes arriving from the primary lodgement, or in phagocytes derived from the blood vessels or tissues of the node itself as a result of local inflammation; and thirdly, that in acute experimental infections the bacteria so held disappear from the nodes within 18 hours or so and may be presumed to be destroyed there. The last two facts, firmly established by the recent work of Smith and Wood (1949) with pneumococcal infections of the rat, suggest an effective and valuable second line of defence that comes into play when the primary lodgement is not eliminated.

For the moment we may defer discussing the reasons for this rather cautious conclusion and pass on to the last stage—the systemic reaction.

Systemic Cellular Defences. We may compare the value of this stage with the preceding two—that is, systemic versus regional

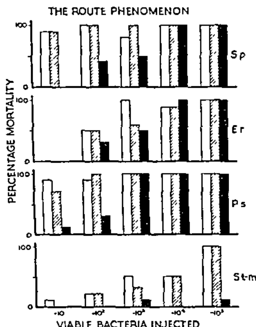


FIG. 7 The death rate in groups of 10 mice given graded doses of various pathogenic bacteria by each of three routes. White column: Mortality by subcutaneous route Shaded (or middle) column: Mortality by intraperitoneal route Black column: Mortality by intravenous route.

S p. *Strep pneumoniae* P.s. *Pasteurella septica*
E.r *Erysipelothrix rhusiopathiae* S.t-m. *Salmonella typhi-murium*

defence—by measuring doses of bacteria that produce fatal infection by the intravenous and, say, by the subcutaneous route. With some pathogens the result is quite at variance with the common mode of thought about the body's first and second line of defence, because the intravenous killing dose is larger than the subcutaneous. Figure 7 summarizes some titrations by

Dr. A. C. Dutton (unpublished) of mouse pathogens; Type I pneumococci, *Erysipelothrix rhusiopathiae*, *Pasteurella septica*, and *Salmonella typhi-murium*. The white columns represent death rate after subcutaneous and the black columns death rate after intravenous injection. Now in these four infections, it is clear that cutaneous tissues and regional lymph nodes of the mouse are, relatively speaking, a liability rather than an asset and that it would have been better for the animal if the subcutaneous bacteria had been rapidly removed to the circulation, where a much bigger dose is readily suppressed. Similar results were obtained with *Bact. friedländeri* and *Strep. faecium*. Not all bacteria behave like this. With *B. anthracis* the intraperitoneal dose was most effective, and the subcutaneous and intravenous equally less effective. We may if we wish regard the evidence in Figure 7 of the greater antibacterial efficacy of a reticulo-endothelial clearance compared with the local reaction, simply as a measure of the very high factor of safety characterizing the body's last line of defence. But it is equally legitimate to make it the basis for questioning the value of the *localizing* function of the response to the primary and lymph node lodgements. May not survival be best ensured by a rapid carriage of the invader from the primary lodgement straight to the Kupfer cell; and may not dispersal rather than localization be the better defence reaction?

Dispersal as a Defence Mechanism

The rich plexus of lymphatic channels under the skin and mucous membranes in many animals, invites the speculation that a ready means of conveying invading pathogens to the lymph nodes and beyond, has contributed to the survival of these species in the face of bacterial competition. But we must take care here not to be misled into dubious teleology; because it is possible that these abundant subsurface lymphatics merely reflect the rich and active vascular plexuses needed for the digestive and respiratory tracts, and for heat-regulating organs like the skin; and that their defensive role is secondary and perhaps accidental.

There is another reason for doubting the universal efficacy of

the lymphatic plexuses as channels for the dispersion of invaders Rich though these plexuses are, the site of a small *natural* lodgement of microbes may be a long way from the nearest lymph channel; and experimentally at any rate, bacteria are extremely reluctant to move through tissue spaces, apparently preferring to adhere to cells. This affinity for cells, already established with some viruses, deserves further study among bacteria. It may be specific, in the sense that *Haemophilus pertussis* is specifically adsorbed to certain red cells—or it may be a consequence of some general property of the bacterium, like surface charge. Nearly a quarter of a century ago Falk and others (see Falk, 1928) demonstrated a correlation between surface charge, which would in part determine affinity for tissue cells, and virulence among the pneumococci. Indeed, the reputed modifications of virulence by surface active agents may be due to non-specific changes of this kind in the affinity of bacteria for tissues. I mention these facts as reasons for regarding the localization in a primary lodgement, not as the achievement of an exquisitely regulated defence mechanism, but merely as the consequence of deplorable bacterial qualities. The bacterium may dictate the site of the primary battleground, by sticking to it; and the phagocytes may consequently gather there because they have no choice.

It is conceivably the function of the phagocyte to aid dispersion as well as to digest the invader. The observation that phagocytes do not necessarily kill the microbes they ingest is an old one, and is usually cited as an example of a failure in tissue defence—a means, for example, whereby the phagocyte-borne microbe is protected until it settles in other tissues to produce a metastatic abscess. The orthodox view of the phagocytes as a carrier of microbes from one part of the body to another is naturally unfavourable, because the observation is usually made on those occasions when such carrying has resulted in an extension of the disease. But a mechanism whereby a microbe that tends to stick to the tissues of an animal could be trans-

have very great survival value in infections demanding defence by dispersion.

These are matters for experiment. I stress them because they raise the important question of interpretation. It is true, as Menkin describes, that a localization of the invader is associated with intercellular fibrin and lymphatic thrombosis in infected tissues; but it commonly occurs only in tissues so heavily damaged that a coagulative necrosis is induced—conditions that may be quite irrelevant to the subtle march of an invader from a clinically undetectable primary lodgement. Again, Smith and Wood's admirable studies of the reactions of infected lymph nodes were made with a dose of pneumococci that killed the animal in 48–72 hours; and the lymph nodes reacted in vain. In the apparent suppression of the pneumococci in the regional node after 24 hours, do we witness a noble failure to stem an irresistible tide, or the irrelevant antics of a tissue overstimulated by a massive infection of a kind unlikely to occur outside the laboratory?

Conclusion

Our intrapolations from the grosser experimental procedures to the delicately balanced adaptations which decide the course of natural infection may be correct; but we need far more evidence before we can accept them as the final answer. We can probably extract the required evidence from these complex non-specific reactions, but only by a gentle analysis into elements with characteristic peculiarities; elements in fact, with characters specific enough to let us recognize their precise role in the more general reaction.

I have indicated the lines on which this detailed analysis might proceed. It may be felt that so far little more has been done than to confirm the position of the microphage plus lymphoid-macrophage system as the chief and most resilient participant in specific and non-specific defence against microbial invasion; an adaptive system which, with the outstanding exception of specific immunization, we have hardly exploited at all. But at least its proved resistance to all but the savage pickaxe-and-dynamite assaults made on it with grams of corti-

sone or hundreds of roentgens of X-irradiation has revealed this system as one of the more enduring rocky slabs in the buried foundation of medicine; and there is no reason why, particularly if as experimentalists we exchange the pickaxe for the geologist's hammer or even the archaeologist's camel-hair brush, we should not soon decipher on it some useful inscriptions.

REFERENCES

BJORNESJO, K. B (1952). *Acta tuberc. Scand.* **27**, 116, 123, 134

BIOZZI, G., BERACERRAF, B. and HALPERN, B (1953). *Brit J. exp. Path* **34**, 441.

BLOOM, W. L, HUDGINS, P C and CUMMINGS, M. M. (1953). *J. infect. Dis.* **92**, 70.

BLOOM, W. L. and PRIGMORE, J. L (1952) *J. Bact.* **64**, 855.

BLOOM, W. L, WATSON, D W, CROMARTIE, W. J. and FREED, F. (1947). *J. infect. Dis.* **80**, 41

BRAUN, W. (1949). *J Bact* **58**, 291.

COBBETT, L. and MELSOME, W S. (1898). *Zbl allg Path path. Anat.* **9**, 827.

DUBOS, R. J. (1950) *J exp Med.* **92**, 319

EGOROFF, A. and LAPTEWA-POPOWA, M (1935) *Acta med Scand.* **87**, 345.

EVANS, D G MILES, A A and NIVEN, J S F. (1948). *Brit. J. exp Path* **29**, 20.

FALK, I S (1928). *The Newer Knowledge of Bacteriology and Immunology*, Chicago University Press, p. 565

FRANCIS, T. (1950) *J Immunol* **65**, 437

[illegible]

therapy,

[illegible]

MAALØE, O (1948) *Acta path microbiol Scand.* **25**, 237.

MCMASTER, P D. and HUDACK, S S (1935). *J exp Med* **61**, 783

MACKAY, M E, MILES, A A., SCHACHTER, M and WILHELM, D. L. (1953) *Nature, Lond* **172**, 714

MENKIN, [illegible]

MENKIN, [illegible]

MENKIN, [illegible]

MICKS, D [illegible] *Biol, N* [illegible] 70, 513.

S

MILES, A. A. and NIVEN, J. S. F. (1950). *Brit. J. exp. Path.* **31**, 73.
MYRVIK, Q. and WEISER, R. S. (1951). *Amer. Rev. Tuberc.* **64**, 699.
1, 247.
J. Path. Bact. **63**, 33
90, 555, 567.
Res Mem.
No. 29
TALIAFERRO, W. H. and TALIAFERRO, L. G. (1951). *J. Immunol.* **66**, 181.
J. Path. **29**, 217.
WEIR, J. A., COOPER, R. H. and CLARK, R. D. (1953). *Science*, **117**, 328.

XV

Antiviral Immunity

H. B. MAITLAND

The subject of antiviral immunity falls naturally for purposes of discussion into two main topics, namely, the basis of antiviral immunity in general, and the means of immunizing against particular diseases. In dealing with each of these aspects the intention is to integrate some of the factors concerned so as to outline the subject as a whole and to indicate points at which the subject is growing rather than to deal exhaustively with any one part of it.

At the beginning the question may be asked, what is the essential nature of antiviral immunity? Is it chiefly humoral or chiefly cellular? What mechanisms are involved? Are these similar to bacterial immunity? The answers are—antiviral immunity has both humoral and cellular factors; in some respects antiviral immunity resembles antibacterial immunity, but there are differences due to the parasitism of viruses and to special phenomena arising from the interaction of virus and host cell.

The Humoral Factor in Immunity

There is no doubt about the importance of antibody in antiviral immunity. This is indicated by the specificity of immunity, passive immunization, the neutralization test in which antibody mixed with virus *in vitro* neutralizes its infectivity, the absorption of antibody by elementary bodies and the fact that some serological tests indicate immunological specificity. All this is self-evident and need not be elaborated.

Viral Antigens

There are, however, some features of viral antigens which deserve mention. The elementary body is complex morphologically as revealed by electron microscopy; it is also complex chemically and antigenically. Analysis has enabled a number of separate antigenic fractions to be specified for some viruses. Where this has been done the evidence indicates or strongly suggests that only one of these antigens is concerned in immunity. It may be assumed that a similar situation exists with other viruses which have not yet been subjected to antigenic analysis, and that in this respect viruses in general accord with bacteria where, as is well known, only one of a number of antigens possessed by a bacterial cell may be concerned with immunity.

The immunizing viral antigen is revealed by the neutralization test and by passive immunization, and these tests indicate therefore immunological similarity or difference in viruses. Other serological tests may reveal other antigens not concerned in immunity which may be shared by related strains or species and give cross-reactions. Thus, depending on which antigenic fraction is operative in complement-fixation, haemagglutinin inhibition or agglutination, these tests may or may not show the same antigenic relationship between viruses as the neutralization test.

Some virus antigens become detached from the elementary body and appear as soluble antigens. In this state they can take part in serological tests, particularly complement-fixation. The immunizing antigen may occur in soluble form but more often it is other antigens which do so. In illustration of these points a few examples may be mentioned.

The vaccinia virus, for instance, has at least four recognized antigens. The so-called LS antigen is a complex of two components, heat-labile and heat-stable respectively, and there is a separate antibody for each component. This antigen is part of the elementary body but is found too in soluble form in infected tissues. It has been purified and is a protein but in its pure form it neither immunizes nor absorbs neutralizing antibody. Another entity is a nucleo-protein antigen which forms a large

part of the elementary body but is not concerned in immunity. The haemagglutinin is a small particle occurring as a soluble antigen distinct from the others and it also takes no part in immunity. The immunizing antigen is certainly contained in the elementary body; its exact identity is not known. It is not the purified LS antigen but this antigen may be easily altered so that during purification it becomes ineffective.

The psittacosis virus is known to possess two antigens. One is specific for this virus, is heat-labile and probably a protein; the second is heat-stable, probably carbohydrate and common to all the viruses in the psittacosis-lymphogranuloma group. The specific antigen is probably the immunizing one and also the toxic factor. Both antigens are soluble, give a complement-fixation test and an allergic skin test.

The influenza virus has a complex immunizing antigen which is part of the elementary body and capable of undergoing alteration in its specific character, thus giving rise to strain differences of immunological importance; it contains a component derived from the host. This antigen is also the haemagglutinin. In addition there is a soluble antigen which is less specific, which does not reveal strain differences but distinguishes between types of influenza virus and is readily detected by complement-fixation. The antigenic relationship of strains revealed by complement-fixation will therefore depend upon which antigen is being used for the test.

Similar though perhaps incomplete data have been obtained on some other viruses but for many exact information about antigenic structure is not available. Immunological relationships have been worked out in the first instance on the basis of cross-immunity experiments and the neutralization test. Serological reactions, often complement-fixation, have then been tried to determine whether they would reveal a similar pattern of antigenic relationships. In some cases, such as in foot and mouth disease, complement-fixation has proved of great value in identifying and differentiating types of virus, undoubtedly because, empirically, the immunizing antigen was used in the test.

In the course of time, much more will be learned about

the antigens of viruses hitherto not studied in detail. The tendency is to place most importance on the immunizing antigen, because of practical considerations related to the control of infectious diseases, and therefore to use it for purposes of classification.

Neutralizing Antibody

One of the central questions of antiviral immunity is, how does antibody neutralize the infectivity of virus? There are three elements in this process, the host cell, the virus and the antibody. How then do these three react? The antibody, as would be expected, combines specifically with the antigen of the elementary body. A purified suspension of elementary bodies will absorb neutralizing antibody from serum. The union is not firm at first. Mixtures of virus and antibody which are non-infective by one route of injection or for one species may be infective by another route or for another species. The infectivity of such mixtures has also been restored by dilution. Firm union between virus and antibody may require several hours or even longer. Nevertheless union of antibody with a particular antigen of the elementary body is the basis of neutralization of infectivity.

The mechanism by which neutralization is effected is not clear; there would seem to be at least three possibilities. Firstly the antibody might have a direct destructive action on the virus. There is no evidence, however, that antibody has an effect comparable with bacterial lysis or that the virus is killed by the antibody. Complement is not a necessary adjunct to neutralization by serum *in vitro* and non-infectivity of virus-serum mixtures can be demonstrated by inoculating them into tissue cultures (e.g. poliomyelitis) in which complement is absent, so that it is probably safe to assume, even when such mixtures are inoculated into animals, that neutralization does not depend on complement.

The second possibility is that union of antibody with virus prevents the virus from combining with cells which normally it can infect. This view is commonly considered to be the most likely explanation of how antiviral serum neutralizes; it may be the correct view but there is no direct proof that it is. The evidence is indirect and does not exclude other interpretations.

Briefly stated it is that antibody, to bring about neutralization, must act on the virus before the virus combines with the cell. This has been shown, for example, by injecting antiserum into an area of skin before virus and thus preventing infection; but if virus is injected first, even a few minutes before serum, infection occurs. Similarly it has been shown *in vitro* (Rous, McMaster and Hudack, 1935), using suspensions of rabbit embryo cells and vaccinia virus, that virus which had combined with living cells was not neutralized by antiserum. By contrast, virus combined with killed cells was neutralized, indicating that some property associated with the life of the cell was involved in protecting the virus from the neutralizing action of the antiserum.

Further evidence has been deduced from haemagglutination of fowl red blood cells by influenza virus. To bring about haemagglutination the elementary bodies combine in large numbers with the red blood cells. The attachment of virus to red blood cells and to tissue cells which it can infect is similar so that factors which affect haemagglutination would be expected to apply to infection of cells. As antiserum will prevent haemagglutination it appears to be assumed that the effect of the antibody is to prevent virus attaching itself to cells. The assumption may well be correct; on the other hand virus combined with antibody could conceivably attach itself to a cell but not exert its characteristic effect. It is a point that could be tested by electron-microscopy, not only with influenza virus and red blood cells but possibly with other viruses and tissue cells. I am not aware of any observations of this kind that have been made.

The third possibility is that elementary bodies having been [illegible]

growth might therefore be affected separately, as has been shown to be the case with some chemical agents.

There is no evidence that neutralizing antibody acts on cells. In direct experiments with cell suspensions (Rous, McMaster and Hudack, 1935) it appears not to do so: and on the basis of specificity there is no reason why it should.

Phagocytosis

In considering the fate of neutralized virus the question of phagocytosis arises. Does something comparable with opsonization of bacteria occur with elementary bodies and are they ingested and destroyed by leucocytes or by fixed reticulo-endothelial cells? Comparatively little of a precise nature has been done on this subject. A number of earlier observations, mostly with vaccinia, indicate qualitatively that blood leucocytes can take up virus, and some cells, particularly monocytes, have been known to destroy it.

A more comprehensive study has been made by Meyer (1941) with psittacosis virus. This virus was large enough to be revealed by ordinary microscopy and the virus particles taken up by leucocytes could be counted. Cell suspensions, consisting predominantly of polymorphonuclears or monocytes, were obtained from peritoneal exudates of mice or guinea-pigs after injecting starch or oil. Cell-virus mixtures were rotated in tubes, films made and stained and phagocytosis estimated by direct counting. Infectivity of the mixtures could be determined by injecting mice. The effect of normal and immune serum and complement could be studied.

Although a little phagocytosis occurred with both polymorphonuclears and monocytes in the presence of normal serum or inactivated immune serum, the striking finding was the very marked effect of unheated immune serum on the ingestion of virus by macrophages. Meyer stated that 'even a casual examination of the cells revealed a maximum crowding of the macrophages with deeply stained typical elementary bodies'. This effect was not seen with polymorphonuclears. Virus treated with unheated immune serum and washed was phagocytosed in the absence of the serum. Thus it appears that the virus was opsonized by the immune serum and that macrophages were much more active than polymorphonuclears in ingesting it. There was

evidence too that phagocytosed virus was destroyed, particularly in mixtures of macrophages and unheated immune serum. Leucocytes derived from immune animals showed considerable antiviral activity in presence of normal serum as though the cells themselves possessed some immune body.

To what extent these results have general application is difficult to say. The psittacosis virus morphologically and in size is unlike many other viruses. Nevertheless it would seem probable that this type of investigation could be applied more widely perhaps with the aid of electron-microscopy. It is interesting that Beard and Rous (1938) found that the activity of vaccinia virus was lessened or suppressed when mixed with a suspension of living reticulo-endothelial cells and that the effect of these cells was far greater than that of polymorphonuclear cells. They also noted that the antiviral principle circulating in animals recovered from vaccinia may be carried through washings by blood leucocytes and retains under such circumstances its neutralizing capacity.

A method of experimentation which is currently being carried on in my laboratory is suitable for investigation in this field. We are engaged in studying quantitatively the fate of vaccinia virus in tissue *in vitro*. The rate and amount of absorption of virus can be measured and its subsequent decrease or increase followed. The method is suitable for systems in which the virus will ultimately grow or in those in which it will be destroyed. All the virus in the system can be accounted for and conditions can be varied at will. A chorio-allantoic membrane is put into a known amount of virus for a stated time, removed, ground and titrated. The virus remaining in suspension is also titrated These two amounts are not significantly less than the virus originally in suspension and are often considerably more The reason for this apparent increase of infective units is not known but it may be due to dispersion of virus caused by contact with the membrane or substances derived from it The amount of virus found on the membrane stated as a percentage of the original suspension is very variable but much of it is loosely attached. If the membrane is washed well and the virus which remains firmly attached to the membrane is titrated

then the amount is remarkably uniform and falls within a relatively narrow range of values. This amount may be taken as a base line for following the fate of the virus. Similarly different kinds of tissue or cell suspensions can be used and the effect of different serum components studied. Beard and Rous (1938) noted the apparent anomaly that although vaccinia virus was destroyed by Kupffer cells in saline suspension, tissue cultures of these cells enabled marked growth of the virus to take place. They state that the two observations were diametrically opposed so far as the fate of the virus was concerned. Meyer (1941) noted a similar phenomenon with psittacosis virus. There are good reasons for accepting that the reticulo-endothelial cells of the tissues are sites for growth of vaccinia, psittacosis and some other viruses. Obviously there are unresolved problems which need further investigation. The method outlined above should be capable of making a contribution in this field.

Viral Toxins

Some viruses are known to have a toxic as distinct from an infective property. This has been shown for psittacosis virus and other viruses of that group, influenza virus, Newcastle disease virus, and less certainly for the viruses of mumps and eastern and western equine encephalomyelitis. The toxin has not been separated from the elementary body and very large doses of virus are required to produce the effect. The toxic action has been differentiated from infection on the grounds of a shorter time to death accompanied by lesions which were not typical of infection, and by injecting virus into an organ in which it did not multiply yet caused lesions and death. The toxins are very labile, but influenza virus has been killed by ultra-violet irradiation without destroying its toxicity.

Toxicity can be neutralized by antiserum. In psittacosis virus the so-called L, or labile, antigen may be the toxic entity though this has not been proved; it also appears to be the immunizing antigen. The psittacosis-lymphogranuloma venereum group of viruses has been divided into six categories based on the specificity of neutralization of toxin. Similarly with influenza viruses, neutralization of toxin has been found to be

specific for types A and B and to a certain extent for strains within the types.

With regard to immunity the role of toxin in the production of lesions concomitant with infection has not been assessed, and therefore the importance of its neutralization as a factor in immunity is not known. Viral toxins have not been separated from elementary bodies as the classical bacterial toxins have been separated from bacteria and nothing comparable with bacterial antitoxic immunity has been envisaged for viruses. The toxin appears to reside in one of the antigens of the elementary body. Whether this antigen is the immunizing antigen, as the specificity of neutralization might suggest, is not certain; in psittacosis it may well be; in influenza it is more doubtful; there need be no uniformity in this respect. In any case these considerations apply at present to a few viruses only.

Allergy

The existence of allergy in virus diseases is not in doubt. It can be demonstrated for a number of viruses and probably occurs to a greater or less extent in virus diseases generally. Although it is due to an antigen-antibody reaction, antigens and antibodies other than those concerned in immunity may be responsible. There is probably no direct relationship between [illegible]

and intensity of the tissue reaction, it is an element in immunity. The manifestation of allergy in virus diseases may be of two kinds, a skin reaction and an accelerated response to infection.

The skin reaction is typified by the Frei test in lymphogranuloma venereum. It is elicited by injecting heated materials containing the so-called S, or heat-stable, antigen. This antigen is thought not to be the immunizing antigen and it is shared by other viruses of this group. The Frei test therefore is not specific for lymphogranuloma—it may be positive in psittacosis—nor is it an indication of specific immunity.

A similar skin reaction has recently been demonstrated with the specific antigen of psittacosis (the L or heat-labile antigen),

which is probably the immunizing antigen. *This reaction is* given only by cases of psittacosis. Probably a comparable result could be obtained with a specific antigenic fraction of the lymphogranuloma venereum virus.

In vaccinia, herpes, influenza and *mumps skin* reactions indicating allergy have been described. In herpes and influenza this occurrence does not coincide with *the presence* of neutralizing antibody in the blood. In mumps, however, those who react are *said to be less liable to* infection but the correlation is far from complete. The immunological significance of the skin reactions to these three viruses is not well defined. The antigenic fraction concerned in each case *has not been adequately* characterized.

In vaccination the introduction *of living virus* into the skin of *persons* who *have not previously been vaccinated causes a* 'normal' reaction which reaches its peak in 8–12 days; those who have been vaccinated before may have an 'accelerated' reaction *with a peak at 4–7 days*, which is regarded as a sign of allergy; the so-called 'immune' reaction is minimal and reaches its peak in 2–3 days. The last two responses have to be interpreted in terms of immunity and allergy.

Heat-killed washed elementary bodies injected intradermally will cause an *allergic reaction in some* persons who have been vaccinated. In appearance it is very like an 'immune' reaction but it *gives no* indication of the degree of immunity towards *revaccination. The 'immune' reaction after vaccination* may be regarded as an 'allergic' reaction in an immune or almost [illegible] giving this type of reaction are [illegible]

The accelerated reaction [illegible] nething approaching the normal response after primary *vaccination* to a minimal 'immune' reaction. Throughout this range there is *infection*, therefore susceptibility, but of varying degree. The acceleration may be looked upon as *a sign of allergy. Thus the* 'accelerated' and 'immune' reactions to vaccination are due to a combination of allergy and immunity *operating* concurrently. The two *factors are* not linked and there is no indication that one is a function of the other.

An accelerated reaction is seen also in the generalized skin eruption in mouse-pox, or ectromelia, as described by Fenner (1948) after intradermal inoculation of the virus. The virus multiplies locally, spreads through the blood to internal organs where it continues to multiply, causing a viraemia with deposition of virus in the skin about the sixth day. The resulting skin lesions develop much more rapidly than does the primary lesion at the site of inoculation, thus indicating the development of allergic sensitization. A similar pathogenic sequence probably occurs in natural smallpox and in measles and other exanthemata, although in these diseases the site of the primary multiplication of virus is not evident.

To recapitulate, allergy and immunity may be due to different antigens, though not necessarily so for it seems likely that in some instances the immunizing antigen may induce allergy. Allergy and other immunity mechanisms can operate at the same time and their respective parts in the total response to infection may be difficult to disentangle. Both humoral and cellular factors in antiviral immunity may be conditioned by allergy.

Cellular Immunity

As a matter of convenience cellular immunity may be defined as [illegible]

tests? Or may they be attached to cells and thus not found in the serum? Or are they released locally from lymphocytes or other antibody-producing cells? Such questions need to be considered when a tissue appears to be specifically resistant to infection in the absence of antibody, as is sometimes the case. Some explanation is required. The specificity of the resistance, when it can be shown, tends strongly towards the conception of an antigen-antibody mechanism.

Possibly in these instances we should look for antibodies which react with antigens below the threshold of detection by customary methods of demonstration.

There are, however, two phenomena of interaction between cell and virus, namely interference and haemagglutination,

which although they are clearly independent of antibody, have an intriguing significance in relation to antiviral immunity.

Interference

The 'interference phenomenon' means in essence that one virus by direct contact with cells or tissue of a host so modifies them that a second virus is prevented from producing its characteristic effects. There are many examples, of which the following will serve to illustrate. A neurotropic yellow fever virus has protected against a viscerotropic strain when the two were given simultaneously to monkeys; a non-neurotropic strain of influenza virus in the brain of mice has protected against Western equine encephalomyelitis and other encephalitis viruses; Coxsackie virus has interfered with infections by poliomyelitis virus.

The two viruses may have no antigenic relationship and the effect therefore is not due to antibodies. Interference is exerted almost at once, before antibodies could be formed. The effect has been demonstrated by giving the virus that causes interference after the infecting virus but a much greater dose is then required. Interference may be partial depending on dose and time of administration.

The effect lasts only a limited time, from a few days up to 2–3 weeks, and is thus much shorter in duration than specific active immunity.

It is a local phenomenon and depends upon contact of elementary bodies with cells. It may be limited to one organ such as brain or allantoic cavity, the rest of the animal or embryo being unaffected.

Killed virus has caused interference, in particular influenza virus killed by ultra-violet irradiation or heat. The factor concerned is, however, very labile. But this finding showed that multiplication of virus was not essential.

The action is not universal. A strain of virus may interfere with other strains of the same species or with some other species, but not with all viruses.

The mechanism of this effect is unknown but an explanation would obviously make an important contribution to the understanding of cell-virus interaction.

Interference thus appears to be a true tissue immunity. In

in the body or persist in the tissues, might exert an interference effect over long periods. There is no certainty that attenuated virus, used for active immunization, does persist indefinitely in tissues but there is evidence that it may persist for a time in some instances. It is possible therefore that interference does play a part in immunity induced with living virus, though to what extent is at present speculative. The duration of active immunity is however much longer than could be accounted for by an interference effect of the inoculum itself, moreover the time of onset of immunity is not like an interference effect. Immunization with dead virus is a different matter and interference is unlikely to operate, for the property of the virus responsible for interference is so labile that it is practically

viruses are known to remain in a symbiotic relationship with tissues for indefinitely long periods, in some cases for life. There are many examples, such as herpes after natural infection in childhood, lymphocytic choriomeningitis after infection of mice *in utero*, or psittacosis or ornithosis in birds. Virus can persist without sign of lesions or illness, but may become active if the balance between virus and tissue is upset. Probably more viruses lie dormant in tissues than are at present recognized, both in man and animals. It is certainly conceivable that they may be exerting an interference effect and thus protecting the cells with which they are associated against pathogenic viruses. Nothing is in fact known of this possibility but it can be envisaged as a natural biological phenomenon of great potential value to the host.

Further work on interference is likely to lead to new facts and conceptions; it is one of the lines of virus research most worth exploring.

which although they are clearly independent of antibody, have an intriguing significance in relation to antiviral immunity.

Interference

The 'interference phenomenon' means in essence that one virus by direct contact with cells or tissue of a host so modifies them that a second virus is prevented from producing its characteristic effects. There are many examples, of which the following will serve to illustrate. A neurotropic yellow fever virus has protected against a viscerotropic strain when the two were given simultaneously to monkeys; a non-neurotropic strain of influenza virus in the brain of mice has protected against Western equine encephalomyelitis and other encephalitis viruses; Coxsackie virus has interfered with infections by poliomyelitis virus.

The two viruses may have no antigenic relationship and the effect therefore is not due to antibodies. Interference is exerted almost at once, before antibodies could be formed. The effect has been demonstrated by giving the virus that causes interference after the infecting virus but a much greater dose is then required. Interference may be partial depending on dose and time of administration.

The effect lasts only a limited time, from a few days up to 2–3 weeks, and is thus much shorter in duration than specific active immunity.

It is a local phenomenon and depends upon contact of elementary bodies with cells. It may be limited to one organ such as brain or allantoic cavity, the rest of the animal or embryo being unaffected.

Killed virus has caused interference, in particular influenza virus killed by ultra-violet irradiation or heat. The factor concerned is, however, very labile. But this finding showed that multiplication of virus was not essential.

The action is not universal. A strain of virus may interfere with other strains of the same species or with some other species, but not with all viruses.

[illegible]

standing of cell-virus interaction.

Interference thus appears to be a true tissue immunity. In trying to assess its possible role in antiviral immunity there are two cases to consider. It might be a factor in immunization with non-pathogenic living viruses. Such a virus, should it multiply in the body or persist in the tissues, might exert an interference effect over long periods. There is no certainty that attenuated virus, used for active immunization, does persist indefinitely in tissues but there is evidence that it may persist for a time in some instances. It is possible therefore that interference does play a part in immunity induced with living virus, though to what extent is at present speculative. The duration of active immunity is however much longer than could be accounted for by an interference effect of the inoculum itself, moreover the time of onset of immunity is not like an interference effect. Immunization with dead virus is a different matter and interference is unlikely to operate, for the property of the virus responsible for interference is so labile that it is practically certain to be destroyed in killing the virus.

Another possibility is that interference may result from living viruses lying dormant in tissues under natural conditions. Some viruses are known to remain in a symbiotic relationship with tissues for indefinitely long periods, in some cases for life. There are many examples, such as herpes after natural infection in childhood, lymphocytic choriomeningitis after infection of mice *in utero*, or psittacosis or ornithosis in birds. Virus can persist without sign of lesions or illness, but may become active if the balance between virus and tissue is upset. Probably more viruses lie dormant in tissues than are at present recognized, both in man and animals. It is certainly conceivable that they may be exerting an interference effect and thus protecting the cells with which they are associated against pathogenic viruses. Nothing is in fact known of this possibility but it can be envisaged as a natural biological phenomenon of great potential value to the host.

Further work on interference is likely to lead to new facts and conceptions, it is one of the lines of virus research most worth exploring.

Haemagglutination

A number of viruses can cause the agglutination of red blood cells but the mechanism by which they do it is not the same for all of them, and only the viruses of influenza, Newcastle disease of fowl, and mumps, concern this discussion. These viruses attach themselves to red blood cells in the same way as they combine with tissue cells before infecting them. What has been learned about haemagglutination has therefore been applied to infection. The facts are well known but may be indicated briefly.

The virus combines with receptors on the red blood cells and appears to act as a union between the cells in bringing them together in clumps. An enzyme of the virus destroys the receptors and the virus is then freed. The red blood cells are no longer agglutinable because they have lost their receptors. A similar enzyme was discovered in cultures of some bacteria, particularly the cholera vibrio; it is known commonly as R.D.E., the receptor-destroying enzyme. Red blood cells treated with R.D.E. are not agglutinable. These findings apply equally to tissue cells—those of the mouse lung and the allantoic cavity of the developing hen's egg. They have similar receptors and union of virus with them appears to be an essential stage in the process by which the virus infects a cell. Their receptors are destroyed by R.D.E. and tissue cells treated by this enzyme are not infected when exposed to virus. This resistance is of relatively short duration. Tissue cells *in vivo* develop new receptors and become susceptible to infection again.

The chief interest in this phenomenon in relation to antiviral immunity is that it affords an instance of cellular resistance due to change in the cell and clearly owes nothing to antibodies. There is no indication that it plays any part in immunity *following infection or active immunization in diseases caused by* these particular viruses. Nor is there any evidence that the [illegible]

as such, it indicates that removal or blockage of cell receptors by whatever means could be a factor in a purely cellular resistance

to viruses. It is not improbable that future investigations with other viruses will show that they too combine with certain substances present in cells, not necessarily in the sense of enzyme and substrate as with influenza virus, and that this general conception may be valid more widely than is now realized.

Immunization against Particular Diseases

The application of general knowledge to the prevention of any virus disease by immunization always raises particular problems which may be discussed briefly on their own account without detailed consideration of any single disease. The antigenic structure of the virus is a primary matter. Strains of a virus may be uniform and stable antigenically as in yellow fever, smallpox, rabies and many other diseases; they may be stable but comprise more than one type, as in poliomyelitis or foot and mouth disease, when it becomes necessary to determine the incidence and distribution of types and select suitable immunizing strains; more difficult still there may be, as in influenza, strain differences which are relatively unstable.

Killed Virus as Vaccine

Another of the main general problems in active immunization is to determine whether it is better to use dead or living virus. The advantage of dead vaccine is its safety; the difficulty is that viral antigens are so labile that their immunizing potency is easily destroyed by the method used to kill the virus. Nevertheless vaccines of killed virus have been prepared which have immunized satisfactorily against infection with the viruses of influenza, equine encephalomyelitis and rabies.

As an example of antigenic lability some early observations with foot and mouth disease virus may be quoted (Maitland, 1928). Virus inactivated by 0·1 per cent formalin at pH 7·6 and 26° C. for 48 hours immunized guinea-pigs against generalized lesions but the antigen was destroyed if the formalin acted for 5 days. Virus kept in phosphate buffer at pH 7·6 and 37° C. for 6–8 days, until it was dead, was a good vaccine but if killed by exposure to pH 6·5 in phosphate buffer for only 29 hours it was useless. A potent formalin-killed vaccine was also made useless

in 24 hours by changing its pH from 7·6 to 6·5. Heating virus at 55° C. and pH 7·6 for 45 minutes also destroyed its antigenicity.

The problem raised by lability applies to viruses generally though some viruses are more labile than others. Of the many agents that have been tried for preparing vaccines of inactivated virus, formalin and ultra-violet light are probably the most generally successful. Perhaps a better way to kill viruses for making vaccines will be discovered. Search in this direction might be rewarding.

One of the most important recent advances in immunization with dead vaccines is the use of an adjuvant to increase their potency. The principle of adjuvants is not new but the type of adjuvant currently employed is a relatively recent introduction. The killed vaccine, usually killed by formalin, is emulsified in a light mineral oil of low viscosity by means of an emulsifying agent so that an emulsion of the water-in-oil type is produced; the virus is contained in small aqueous particles surrounded by an oily matrix (Freund, 1951). This is injected subcutaneously or intramuscularly. The antibody response is many times greater than after vaccine in a saline medium, the titre remains high for a much longer time and the specificity is said to be broader; for example, with influenza the antibody would cover closely related, but not antigenically identical, strains. These advantages have practical significance and may offer new prospects in immunization. Vaccines of this type against influenza and poliomyelitis have been prepared and are undergoing preliminary trials.

Living Virus as Vaccine

Living virus is a much better immunizing agent than killed virus, but there are great difficulties inherent in the use of living vaccine as an immunizing agent. Vaccinia virus which sets up an infection is regarded as an exception to the requirement that living vaccines must be devoid of all harmful effects. That it is possible to obtain strains of virus so modified that they have lost completely their pathogenic properties but retain their immunizing capacity is exemplified by the yellow fever virus. The

production of such non-virulent strains is nevertheless a matter of real difficulty. No standard procedure exists; each virus has to be investigated as a separate problem. Modification is usually attempted by propagating the virus in a species or tissue which it does not normally infect and searching for the desired change in properties. Having obtained a suitable strain it must be thoroughly tested for safety and efficacy in animals and then in man and its continued propagation without reversion of pathogenicity or loss of immunizing potency must be achieved. Technically these problems have been overcome with some viruses and presumably can be for others. Many virologists believe that the efficiency of living vaccines is so superior to dead vaccines that the former are preferable on principle and worth striving for in spite of the difficulties entailed. The idea of using living vaccines against poliomyelitis and influenza is the incentive for much current research. It remains to be seen whether living or dead vaccines will ultimately be preferred for these diseases and for others which await investigation. In some cases it may perhaps be found advantageous to use the two types of vaccine in conjunction; an initial injection of a dead vaccine might be reinforced by a subsequent injection of living vaccine, the increase in resistance arising from the first injection affording an added element of safety in respect of the second.

Pathogenesis and Immunity

The connection between pathogenesis and immunity derives [illegible] ties are most favourable where a viraemia is a natural stage in the development of the infection and least favourable when a virus attacks superficial cells directly.

From this point of view at least three patterns of pathogenesis can be discerned. The first is exemplified by one of the pox diseases. The course of experimental infection has been worked out for mouse-pox (ectromelia) by Fenner (1948). The virus was inoculated into the skin of the foot where it multiplied; in natural infections the primary site of virus growth may be

unknown or surmised from circumstantial evidence. After local growth there follows a transient slight viraemia and settling out of virus in internal organs, probably in reticulo-endothelial cells, where further multiplication takes place, followed by a heavier viraemia with deposition of virus in the skin and the development of skin lesions. The general symptoms of fever, etc., begin at the time of the second viraemia. The virus has to pass through the blood at two stages and in either stage it is exposed to the action of circulating antibody. It is in diseases of this nature that active or passive immunization may be expected to be most effective.

A similar course of events occurs in smallpox. The virus can be isolated from the blood at the time of beginning of symptoms, i.e. during the secondary viraemia. Measles and chicken-pox probably have a similar pathogenesis; the incubation periods are long enough for a succession of stages to occur. Measles has been transmitted to human volunteers and monkeys with blood from early cases. Both measles and chicken-pox can be prevented by passive immunization during the first few days of the incubation period.

A variation of this type of pathogenesis occurs when there is a primary site of multiplication of virus followed by spread through the blood to an internal organ where the main and characteristic signs of the disease develop. It is of great interest and importance that poliomyelitis has recently been shown to have a stage of viraemia and probably conforms to this pattern of pathogenesis. The current view is that the site of primary multiplication is the intestine and that from there the virus spreads by the blood to the central nervous system. That being so there are rational grounds for expecting immunization to be effective in preventing the paralytic stages of the disease, as indeed recent work has shown to be the case in limited preliminary trials.

In diseases which have this pattern of pathogenesis 'prevention' means suppression of the clinical and pathological signs of infection; it does not necessarily mean abolition of what goes on during the early part of the incubation period before the clinical signs become manifest. Thus 'immunity' may not pre-

vent lodgement of virus, primary multiplication and even some degree of spread. A much higher grade of resistance would be required to abolish infection completely. In poliomyelitis, for example, a degree of immunity that would prevent paralysis and be clinically effective might not prevent growth of virus in the intestine.

A second pattern of pathogenesis is afforded by those diseases in which infection is acquired through the skin—often by the bite of an insect—and the virus has to be transported by the blood to the sites where it can grow and produce its characteristic effects. Such for instance is yellow fever. Here too the operation of antibodies is favoured and immunization in advance of acquiring the infection would be expected to be successful, as indeed it is.

In contrast with these diseases there is a third type of pathogenesis in which superficial tissues are primarily attacked, the infection develops fully at the point of entry and there is no secondary spread through the blood. Influenza is an example of this type of disease. Antibodies in the blood are here at a disadvantage; to be effective they have to reach the epithelial surface of the respiratory tract and it is estimated that the strength of antibody in this situation is only one-tenth to one-twentieth of that in the blood. Nevertheless it has been possible to immunize actively against influenza.

Some diseases such as herpes zoster or rabies do not fit well into any of the pathogenic patterns that have been mentioned, but pathogenesis, whatever the pattern, may be regarded as conditioning the operation of the processes of immunity. The length of the incubation period which is one aspect of pathogenesis has been suggested as having a bearing on the duration of immunity following an attack of disease. If there is a long incubation, and especially if this is coupled with a phase of viraemia as it often is, the situation is favourable for a subsequent infection to be headed off and overcome even if no antibody is present in the blood at the start; there would be time for a secondary type of response with its characteristic rapid development of antibodies to take place. Thus immunity would be long-lasting and such aborted infections would themselves

unknown or surmised from circumstantial evidence. After local growth there follows a transient slight viraemia and settling out of virus in internal organs, probably in reticulo-endothelial cells, where further multiplication takes place, followed by a heavier viraemia with deposition of virus in the skin and the development of skin lesions. The general symptoms of fever, etc., begin at the time of the second viraemia. The virus has to pass through the blood at two stages and in either stage it is exposed to the action of circulating antibody. It is in diseases of this nature that active or passive immunization may be expected to be most effective.

A similar course of events occurs in smallpox. The virus can be isolated from the blood at the time of beginning of symptoms, i.e. during the secondary viraemia. Measles and chicken-pox probably have a similar pathogenesis; the incubation periods are long enough for a succession of stages to occur. Measles has been transmitted to human volunteers and monkeys with blood from early cases. Both measles and chicken-pox can be prevented by passive immunization during the first few days of the incubation period.

A variation of this type of pathogenesis occurs when there is a primary site of multiplication of virus followed by spread through the blood to an internal organ where the main and characteristic signs of the disease develop. It is of great interest and importance that poliomyelitis has recently been shown to have a stage of viraemia and probably conforms to this pattern of pathogenesis. The current view is that the site of primary multiplication is the intestine and that from there the virus spreads by the blood to the central nervous system. That being so there are rational grounds for expecting immunization to be effective in preventing the paralytic stages of the disease, as indeed recent work has shown to be the case in limited preliminary trials.

In diseases which have this pattern of pathogenesis 'prevention' means suppression of the clinical and pathological signs of infection; it does not necessarily mean abolition of what goes on during the early part of the incubation period before the clinical signs become manifest. Thus 'immunity' may not pre-

XVI

The Action of Bacterial Enzymes on Immunizing Antigens

LORD STAMP

During recent years the importance of immunological methods in the treatment of bacterial infections has been greatly reduced, owing to the highly successful results that have been achieved with chemotherapy and antibiotics. The use of vaccines and antibacterial sera in treatment has therefore been almost completely abandoned. It seems possible, however, that they may yet have an important part to play in the treatment of some infections at least, in view of certain limitations and disadvantages of chemotherapy, to which Professor Garrod refers in his lecture on 'Causes of Failure in Antibiotic Therapy'.[1] I might mention, for example, the increasing frequency with which drug-resistant strains of some organisms are being encountered, and the tendency for various toxic and allergic manifestations to develop following the repeated use of some drugs, which may become more and more serious problems.

Apart from the possible future need for antibacterial immunization in treatment, it must continue to be a bulwark in preventive medicine, particularly against some infectious diseases. The degree of protection afforded, however, by present-day methods in many cases leaves much to be desired. For example, there is little experimental evidence that staphylococcal vaccines are capable of producing an effective antibacterial immunity. Clinically also, they have given very variable results

[1] See pages 302–13

serve to maintain the basal immunity. In a disease like influenza with short incubation period, and no spread through the blood, there would be neither time for antibody to develop nor the same opportunity for it to act. Thus the secondary antibody response in face of a new infection would not function and there would be relatively short immunity after recovery. This suggestion is no doubt an over-simplification and cannot in itself account for all instances of either prolonged or relatively short immunity after an attack of infection, but on the other hand it is a factor to be considered in the relation of pathogenesis to immunity.

REFERENCES

BEARD, J. W. and ROUS, P. (1938). *J. exper. Med* **67**, 883.
FENNER, F. (1948). *Lancet*, **ii**, 915.
FREUND, J. (1951). *Amer. J. clin Path.* **21**, 645.
MAITLAND, H. B. (1928) '3rd Progress Report of the Foot and Mouth Disease Committee', H.M. Stat. Off, London, p. 41.
MEYER, K. F. (1941). *Schweiz. med. Wschr.* **71**, 436.
ROUS, P., MCMASTER, P. D. and HUDACK, S. S (1935). *J. exper. Med.* **61**, 657.

full activity of the enzyme. This was confirmed by experiments in which extracts of the cells containing the 'M' antigen were treated with partially purified enzyme preparations in the presence of various reducing substances. Under such conditions the 'M' antigen was readily destroyed.

I have obtained evidence of the immunizing antigen-destroying activity of streptococcal proteinase by means of active immunization experiments in mice (Stamp, 1953). Treatment of heat-killed vaccines with a proteinase-containing filtrate at varying pH in the presence of Na thioglycollate was found to result in complete destruction of immunizing activity at both pH 6·0 and 7·0. No such effect was obtained with filtrates previously heated to 70° C. for one hour to inactivate the enzyme.

The proteolytic activity of culture filtrates and partially purified enzyme preparations has been demonstrated on various protein substrates such as milk (Elliott and Dole, 1947; Stamp, 1953), casein, fibrin and gelatin (Elliott, 1945), crude muscle protein, and also a preparation of protein derived from the skin originally described by Oakley, Warrack and Van Heyningen (1946), known as 'azocoll' (Todd, 1947). Elliott has also studied the ability of partially purified proteinase to hydrolyse two synthetic polypeptide substrates, namely α-benzoyl-l-argenineamide and l-leucyl-glycylglycine, a test originated by Bergmann to classify different groups of proteinase (Bergmann, 1942). The hydrolysis of the former but not the latter in the presence of cysteine afforded evidence that the enzyme is similar to papain and some of the cathepsins.

The activity of the enzyme is greatly enhanced by reducing agents such as glutathione, cysteine, thioglycollic acid and also by potassium cyanide, but is inhibited by iodoacetic acid, suggesting that it depends on the reduction to sulphydryl groups of disulphide groups present in the enzyme (Elliott, 1945).

It would appear that streptococcal proteinase may be identical with a muscle-digesting enzyme named 'histase' described by Frobisher (1926). The general proteolytic activity of this enzyme was not recognized at the time probably owing to the

in the treatment of chronic staphylococcal infections. Furthermore, immunization with toxoid, which in pre-chemotherapy days had largely superseded vaccine therapy, though of greater value, is not effective against all forms of staphylococcal infection. Improvements in methods of immunization against staphylococcal infections are all the more necessary, in view of the fact that in these infections in particular, the use of antibiotics is becoming increasingly restricted. To take another example, numerous reports indicate that T.A.B. vaccines prepared by the latest methods afford a very incomplete protection against typhoid fever, particularly in highly endemic areas (Miller *et al.*, 1951). It is evident that in general, much still remains to be learnt of the immunizing antigens of bacteria and of the conditions necessary for their development and preservation in vaccine or culture filtrate.

In recent years evidence has been forthcoming that the failure of killed vaccines or culture filtrates to immunize in some cases at least, is due to the destruction of the immunizing antigen by enzymes produced by the organisms when grown *in vitro*. It is this evidence that I wish to discuss in the present lecture.

The species that has been studied most fully in this respect is *Str. pyogenes*. While antibacterial immunization against this organism is not generally considered practicable or necessary, nevertheless, these studies have produced results of great theoretical interest. Moreover the methods used may well prove of considerable value when applied to similar investigations on other organisms of greater practical importance.

Streptococcal Proteinase and the 'M' Antigen

In 1945, Elliott (1945) demonstrated the presence of a proteolytic enzyme in culture filtrates of certain strains of *Str. pyogenes*. He showed by means of a precipitin test that this enzyme was able to destroy the 'M' immunizing protein of *Str. pyogenes* (Hirst and Lancefield, 1939; Lancefield, 1940; Stamp and Hendry, 1937), particularly when the antigen was present in living cells, but it was much less active after the antigen had been extracted from the cells It was concluded that the reducing conditions produced by living cells were necessary for

full activity of the enzyme. This was confirmed by experiments in which extracts of the cells containing the 'M' antigen were treated with partially purified enzyme preparations in the presence of various reducing substances. Under such conditions the 'M' antigen was readily destroyed.

I have obtained evidence of the immunizing antigen-destroying activity of streptococcal proteinase by means of active immunization experiments in mice (Stamp, 1953). Treatment of heat-killed vaccines with a proteinase-containing filtrate at varying pH in the presence of Na thioglycollate was found to result in complete destruction of immunizing activity at both pH 6·0 and 7·0. No such effect was obtained with filtrates previously heated to 70° C. for one hour to inactivate the enzyme.

The proteolytic activity of culture filtrates and partially purified enzyme preparations has been demonstrated on various protein substrates such as milk (Elliott and Dole, 1947; Stamp, 1953), casein, fibrin and gelatin (Elliott, 1945), crude muscle protein, and also a preparation of protein derived from the skin originally described by Oakley, Warrack and Van Heyningen (1946), known as 'azocoll' (Todd, 1947). Elliott has also studied the ability of partially purified proteinase to hydrolyse two synthetic polypeptide substrates, namely α-benzoyl-l-argenineamide and l-leucyl-glycylglycine, a test originated by Bergmann to classify different groups of proteinase (Bergmann, 1942). The hydrolysis of the former but not the latter in the presence of cysteine afforded evidence that the enzyme is similar to papain and some of the cathepsins.

The activity of the enzyme is greatly enhanced by reducing agents such as glutathione, cysteine, thioglycollic acid and also by potassium cyanide, but is inhibited by iodoacetic acid, suggesting that it depends on the reduction to sulphydryl groups of disulphide groups present in the enzyme (Elliott, 1945).

It would appear that streptococcal proteinase may be identical with a muscle-digesting enzyme named 'histase' described by Frobisher (1926). The general proteolytic activity of this enzyme was not recognized at the time probably owing to the

fact that the reducing conditions necessary for activation were not present in other substrates tested (Todd, 1947).

Streptococcal proteinase appears to be distinct from the fibrinolysin of Tillett and Garner (1933) since according to Elliott, proteinase is able to digest this enzyme, and moreover the two differ considerably in heat resistance.

Elliott (Elliott and Dole, 1947; Elliott, 1950) has also shown that proteinase is produced initially in the form of an inactive precursor, which is activated autocatalytically or by treatment with trypsin. Autocatalytic conversion of precursor to proteinase is initiated by traces of active proteinase and is accelerated by reducing conditions. Relatively pure preparations of precursor and proteinase have been obtained in crystalline form by repeated recrystallization from $(NH_4)_2SO_4$ solutions of varying concentrations, and at different pH and temperatures, *and* furthermore, the two substances have been found to be antigenically distinct.

The proteolytic activity of a culture filtrate depends on a number of factors. In the first place, the strain used is of importance. Todd (1947) tested the proteinase-producing ability of 42 strains of Group A Streptococci. Culture filtrates of 30 gave titres varying from $\frac{1}{2}$ to $\frac{1}{500}$. The remaining 12 were negative. Some of these, however, gave weakly positive results on concentrating the filtrate with $(NH_4)_2SO_4$. It seems probable, therefore, that the great majority of strains are able to produce proteinase and that differences in this respect are nearly always quantitative rather than qualitative.

Proteolytic activity also depends on the degree to which precursor has been converted into active proteinase during growth or on storage, and this is influenced by the medium used, the precise conditions of growth and on the time interval between filtration and testing. Conversion is accelerated in media containing traces of trypsin (Elliott, 1950), or in culture media which maintain relatively low O/R potentials during growth due to low oxygen-carrying capacity and possibly the presence of traces of reducing substances. Filtrates prepared under such conditions have relatively high titres when tested immediately after filtration and little or no increase occurs on storage.

Conversion is retarded at a temperature of 22° C. (Elliott, 1945), or by growing under the relatively high O/R potentials produced by aerating the culture or by using media with a high oxygen-carrying capacity, e.g. certain batches of digest broth (Hewitt, 1950) put up in a shallow layer and not recently autoclaved.

Filtrates prepared under such conditions often have relatively low titres when fresh, but a very marked increase occurs on storage, particularly if they are kept at an acid pH and if Na thioglycollate is added. In one experiment, for example, the milk-clotting titre of the fresh filtrate at pH 8·0 was $\frac{1}{8}$. After adding Na thioglycollate to 0·1 M and storing at 20° C. for 3 days at pH 6·2, the titre at pH 8·0 had increased to $\frac{1}{1024}$, but had increased only slightly, namely to $\frac{1}{16}$, when stored at pH 7·7 under the same conditions (Stamp, 1953). Apart from the factors mentioned above, it is probable that traces of metals may also accelerate or retard the autocatalytic process.

The proteolytic activity of a culture filtrate also depends on the pH developed during growth. When culture filtrates are prepared in media maintaining a relatively high pH during growth, i.e. *circa* 7·0, they show no proteolytic activity even though immunological tests of the heat-killed vaccines prepared from such cultures suggest that proteinase has been produced.

A low final pH, i.e. *circa* 6·0, such as is produced in some batches of digest broth or by the addition of glucose, appears to be necessary for the liberation of precursor or proteinase from the cells and their appearance in the culture filtrate.

The activity of proteinase once it has been formed also depends on the pH at which it is allowed to act. The optimal pH appears to vary according to the method of testing (Elliott, 1945; Stamp, 1953). It would appear, however, that there is little difference in the degree of activity within the range of pH normally developed in streptococcal cultures.

Filtrates derived from cultures grown in different media vary considerably in proteolytic activity, however, even though the degree of growth is the same and optimal pH and O/R potential conditions for the production of active and filtrable proteinase have been maintained. This suggests that there may be factors

in the medium not yet determined which influence the production of precursor.

I have dealt at some length with the factors influencing the production of active proteinase since they largely determine the degree to which the 'M' antigen is destroyed during growth and therefore the immunizing activity of the heat-killed vaccine.

It seems probable that destruction of the immunizing antigen by proteinase bound to the cells may continue, though more slowly, during storage, particularly with strongly proteinase-producing strains, since the heat treatment employed to kill the vaccine, namely 55° C. for 1 hour, is insufficient to inactivate the enzyme.

In the light of these observations, we may now attempt to interpret the results of active immunization experiments in which vaccines of a weakly proteinase-producing strain 'Richards', grown in different batches of digest broth and at *varying O/R potential, were tested for immunizing activity* in mice against the homologous strain. Each vaccine was tested by immunizing batches of 30 mice, as far as possible under the same conditions, the mice receiving the same immunizing and challenge doses. In some series, the percentage survivals recorded *represent the average of several experiments with* vaccines prepared under the same conditions.

TABLE 1 The Immunizing Activity of Vaccines of *Str. pyogenes* Strain 'Richards' Grown in Different Batches of Digest Broth at Varying O/R Potential (Adapted from Stamp, 1953)

Batch of digest broth	Final pH	Proteinase activity in culture filtrate of strain		Percentage survivors in mice immunized with vaccines grown at			
				low Eh	normal Eh	high Eh	Controls
		Richards	William	+ thio or ascorbate	Static	Shaken	
A	6·0–6·8	±	+ + +	0 0	21·7	56 7	1·7
B	7 2	–	–	10 0	10 3	52·5	0 0
C	6·4	–	+	42 0	66 6	–	3·3
D	7·1	–	–	75 0	73 0	80 0	1 7

It will be seen firstly that with Batch A digest broth, which strongly favours proteinase production, the O/R potential exerts a marked influence on immunizing activity. The immunizing activity of vaccines grown at a low O/R potential in Na ascorbate or thioglycollate broth is nil, but when they are grown at high O/R potentials produced by shaking the culture, it is relatively high. The same effect, though less marked, is seen with vaccines grown in Batch B, though owing to the higher final pH, no proteinase was demonstrated in the filtrate even with a strongly producing strain. In Batch C which favours proteinase production to a much lesser extent, relatively highly immunizing vaccines are obtained even when grown at a very low potential. With Batch D this is even more evident, though again absence of proteinase production could not be confirmed, owing to the neutral final pH of the filtrate.

On the other hand, with strongly proteinase-producing strains, factors such as the batch of medium used and the O/R potential developed have little or no influence on immunizing activity, vaccines in general being entirely or almost entirely inactive, even when grown under conditions unfavourable to proteinase production or activation. As already suggested, this may possibly be due to conversion of precursor into proteinase and destruction of the immunizing antigen during storage. Altogether, I have prepared vaccines from nine different mouse virulent strains in media not favouring proteinase production The vaccines of three weakly proteinase-producing strains were actively immunizing, while those of five out of six strongly proteinase-producing strains showed little or no activity.

The experimental evidence taken as a whole therefore strongly supports the view that proteinase activity plays a predominant role in regulating the immunizing potency of vaccines of mouse virulent strains of *Str. pyogenes*, even though certain anomalies still remain to be explained and some gaps in the chain of evidence have still to be filled in.

Anthrax Proteinase and Immunizing Antigen

Another organism which produces an enzyme apparently able to destroy the immunizing antigen is the anthrax bacillus.

Many years ago Bail (1904) showed that the oedema fluid of cutaneous anthrax lesions contained a filtrable antigen capable of immunizing animals against experimental anthrax infection. Bail's observations have since been amply confirmed and extended by many workers (Graber and Staub, 1946; Cromartie, Watson, Bloom and Heckly, 1947). Attempts to produce an immunizing antigen *in vitro* have, however, until recently met with little success. Casagrandi (1902) and Schilling (1927) claimed to have produced some degree of immunity in rabbits by immunizing with filtrates from broth cultures containing high concentrations of serum or plasma. White (1946) has also produced an appreciable degree of immunity in guinea pigs by immunizing with a vaccine consisting of a whole culture of a non-sporing variant of *B. anthracis* grown in whole blood and killed with toluol. In none of these experiments, however, was the immunity induced of a high order. Gladstone (1946), however, has shown that filtrates of cultures grown in serum or plasma under carefully controlled conditions contained an immunizing antigen which was as active as that obtained from oedema fluid. For its production it was necessary that a small inoculum be used, consisting of *circa* 10^4 spores per 25 ml. of plasma, and also that the incubation period should be short, i.e. about 18 hours. After that period the antigen was rapidly destroyed and in filtrates from 48-hour cultures little or no immunizing antigen could be detected. Gladstone (1948) later found that the incubation period could be lengthened to 3 days, and filtrates twenty-five times more potent could be obtained if the cultures were dialysed against broth in a special apparatus designed to ensure a continuous supply of oxygen and fresh nutrients.

More recently, potent filtrates have been obtained in a medium containing hydrolysed casein, yeast extract and charcoal (Belton and Strange, 1954) and in chemically defined non-protein medium (Wright, Hedberg and Slein, 1954). It is evident, therefore, that contrary to the indications of earlier work, high concentrations of serum or plasma are not essential for the production of immunizing antigen.

There are strong grounds for believing, as Gladstone has

suggested, that lack of immunizing activity in culture filtrates grown under certain conditions may at least in part be due to destruction of the immunizing antigen as the result of enzyme action. In the first place there is increasing evidence from immunochemical studies of the fractionated filtrate that the immunizing antigen is a protein (Strange and Belton, 1954).

Secondly, it is well known that broth culture filtrates contain an enzyme attacking casein and other proteins as is shown in Table 2. Thirdly, Gladstone has found that highly proteolytic culture filtrates, as well as proteolytic enzymes in general such as trypsin, destroy the immunizing antigen after overnight incubation (personal communication). There is also the evidence obtained from the study of a so-called non-proteolytic variant of the 'Vollum' strain produced by irradiating with ultra-violet light (Wright, Hedberg and Feinberg, 1951). This strain, referred to as NP-A, was assumed to be completely non-proteolytic since it failed to produce a zone of clearing round the colonies on milk agar plates, characteristic of the parent strain. Furthermore filtrates from three-day cultures in yeast extract broth showed no proteinase activity. Active immunization of rabbits with plasma culture filtrates showed that the NP-A strain could be grown for a longer period than the parent strain without loss of immunizing activity. This suggested that the destruction of the immunizing antigen on prolonged culture noted by Gladstone was partly due to proteinase activity, though the fact that the antigen was destroyed eventually even in cultures of the NP-A strain seemed to indicate that other factors were also responsible.

On the other hand it seems possible that these results might be explained solely on a basis of proteinase activity, since I have found that the NP-A strain when grown under particularly favourable conditions does in fact produce some proteinase though not as much as the parent 'Vollum' strain (Table 2).

It would seem, therefore, that the NP-A strain is more correctly referred to as a weakly proteinase-producing strain analogous to those of *Str. pyogenes*. Nevertheless it might be thought that it would prove superior to the parent strain for the production of immunizing antigen. This, however, has not

Many years ago Bail (1904) showed that the oedema fluid of cutaneous anthrax lesions contained a filtrable antigen capable of immunizing animals against experimental anthrax infection. Bail's observations have since been amply confirmed and extended by many workers (Graber and Staub, 1946; Cromartie, Watson, Bloom and Heckly, 1947). Attempts to produce an immunizing antigen *in vitro* have, however, until recently met with little success. Casagrandi (1902) and Schilling (1927) claimed to have produced some degree of immunity in rabbits by immunizing with filtrates from broth cultures containing high concentrations of serum or plasma. White (1946) has also produced an appreciable degree of immunity in guinea pigs by immunizing with a vaccine consisting of a whole culture of a non-sporing variant of *B. anthracis* grown in whole blood and killed with toluol. In none of these experiments, however, was the immunity induced of a high order. Gladstone (1946), however, has shown that filtrates of cultures grown in serum or plasma under carefully controlled conditions contained an immunizing antigen which was as active as that obtained from oedema fluid. For its production it was necessary that a small inoculum be used, consisting of *circa* 10^4 spores per 25 ml. of plasma, and also that the incubation period should be short, i.e. about 18 hours. After that period the antigen was rapidly destroyed and in filtrates from 48-hour cultures little or no immunizing antigen could be detected. Gladstone (1948) later found that the incubation period could be lengthened to 3 days, and filtrates twenty-five times more potent could be obtained if the cultures were dialysed against broth in a special apparatus designed to ensure a continuous supply of oxygen and fresh nutrients.

More recently, potent filtrates have been obtained in a medium containing hydrolysed casein, yeast extract and charcoal (Belton and Strange, 1954) and in chemically defined non-protein medium (Wright, Hedberg and Slein, 1954). It is evident, therefore, that contrary to the indications of earlier work, high concentrations of serum or plasma are not essential for the production of immunizing antigen.

There are strong grounds for believing, as Gladstone has

proved to be the case according to Wright and his colleagues (Wright, Hedberg and Slein, 1954), and Belton and Strange (personal communication), possibly owing to the fact that in the medium used proteinase production was completely inhibited even with the strongly proteinase-producing parent strain. I have, in fact, failed to demonstrate any proteinase in culture filtrates of the 'Vollum' strain grown in hydrolysed casein yeast extract charcoal medium, prepared by Belton.

Further evidence of the destructive action of proteinase on the immunizing antigen might be obtained by establishing a correlation between the presence of proteinase and absence of immunizing antigen and vice versa in filtrates of cultures grown under varying conditions. Such evidence is however very incomplete.

Gladstone (1948) found that, judged by protein estimations, proteolysis was absent in highly immunizing serum filtrates grown under the special conditions referred to above, whereas it was marked in ordinary stagnant cultures, particularly in the later stages of growth when the immunizing antigen was destroyed.

On the other hand, proteolysis was also absent in aerated plasma-culture filtrates showing little or no immunizing activity. It seemed, therefore, that under such conditions other factors such as denaturation of the antigen or inhibition of its production due to the antifoaming substances used might also be responsible for lack of immunizing activity.

I have recently been studying the effect of varying the growth conditions on the production of proteinase, with particular reference to the influence of O/R potential. The results of this investigation, taken in conjunction with the findings of other workers with regard to the conditions required for the production of immunizing antigen, appear to have some bearing on the problem.[1] Cultures of the 'Vollum' strain were grown at 37° in very shallow layers (10^7 spore inoculum/60 ml. of medium

[1] Some of these observations were made after this lecture was delivered.

TABLE 2. The Proteinase Titres with and without Na thioglycollate of Culture Filtrates of *B. anthracis* Grown for Varying periods in Different Media, Tested immediately after Filtration and also after Storage

Strain	Medium	Days stored 18°C.	Clotting of milk						Liquefaction of gelatin					
			Age of culture						Age of culture					
			24 hr.		48 hr.		96 hr.		24 hr.		48 hr.		96 hr.	
			Na thio (0·1 M) substrate						Na thio (0·1 M) substrate					
			+	−	+	−	+	−	+	−	+	−	+	−
Vollum	Horse serum	0	—	—	—	—	—	—	—	—	—	$\frac{1}{64}$	—	$\frac{1}{256}$
		6	—	—	—	—	—	$\frac{1}{2}$	—	$\frac{1}{2}$	—	$\frac{1}{128}$	—	$\frac{1}{512}$
		14	—	—	—	—			—	$\frac{1}{4}$	—	$\frac{1}{256}$		
Vollum	20% horse-serum broth	0	—	—	—	—	—	$\frac{1}{2}$	—	$\frac{1}{4}$	—	$\frac{1}{16}$	—	$\frac{1}{256}$
		6	—	—	—	—	—	$\frac{1}{32}$	—	$\frac{1}{4}$	—	$\frac{1}{256}$	—	$\frac{1}{1024}$
		14	—	—	$\frac{1}{4}$	$\frac{1}{128}$	$\frac{1}{256}$	$\frac{1}{256}$	—	$\frac{1}{8}$	—	$\frac{1}{1024}$	$\frac{1}{32}$	$\frac{1}{2048}$
Vollum	Infusion broth	0	—	$\frac{1}{8}$	—	$\frac{1}{8}$	—	$\frac{1}{32}$	—	$\frac{1}{64}$	—	$\frac{1}{128}$	—	$\frac{1}{512}$
		6	—	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{64}$	—	$\frac{1}{32}$	—	$\frac{1}{256}$	$\frac{1}{4}$	$\frac{1}{2048}$
		14	$\frac{1}{16}$	$\frac{1}{64}$	$\frac{1}{64}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{256}$	$\frac{1}{8}$	$\frac{1}{512}$	$\frac{1}{32}$	$\frac{1}{2048}$	$\frac{1}{16}$	$\frac{1}{2048}$
Vollum	Infusion broth + Na thio (0·01 M)	0	—	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{64}$	—	$\frac{1}{32}$	—	$\frac{1}{128}$	$\frac{1}{8}$	$\frac{1}{512}$	$\frac{1}{2}$	$\frac{1}{512}$
		6	—	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{2}$	$\frac{1}{64}$	—	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{1024}$	$\frac{1}{4}$	$\frac{1}{2048}$
		14	$\frac{1}{8}$	—	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{16}$	$\frac{1}{128}$	—	$\frac{1}{4}$	$\frac{1}{64}$	$\frac{1}{1024}$	$\frac{1}{8}$	$\frac{1}{2048}$
NP-A	Infusion broth	0	—	—	—	$\frac{1}{2}$	—	$\frac{1}{4}$	—	$\frac{1}{2}$	—	$\frac{1}{32}$	$\frac{1}{2}$	$\frac{1}{32}$
		6	—	—	—	—	—	$\frac{1}{2}$	—	—	—	$\frac{1}{32}$	—	$\frac{1}{64}$
		14	—	$\frac{1}{2}$	—	$\frac{1}{4}$	—	$\frac{1}{4}$	—	$\frac{1}{4}$	—	$\frac{1}{64}$	—	$\frac{1}{32}$

— = < $\frac{1}{2}$.

significant that other workers have found this medium to favour the production of immunizing antigen provided incubation is not prolonged. In 20 per cent serum broth, proteinase production though marked after prolonged culture is also delayed, and this may perhaps be related to the delayed fall in O/R potential.

The second point to be noted is that the titre of proteinase varies considerably with the agĕ of the filtrate. Many filtrates show little or no activity when fresh but are markedly proteolytic after keeping for 14 days. This phenomenon is reminiscent of that which occurs with streptococcal proteinase. It suggests that anthrax proteinase like streptococcal proteinase may under certain conditions be produced in the form of an inactive precursor which is activated autocatalytically, though unlike streptococcal proteinase at a relatively high O/R potential. Not all filtrates show this phenomenon however. For example, in the 24-hour Na thioglycollate broth culture filtrates the reverse process occurs, indicating that under certain conditions proteinase may be destroyed These findings, however, show that in any attempt to study the proteinase activity of a culture filtrate in relation to its immunizing activity, it is important that the filtrate should be tested some time after filtration as well as when fresh, since it is quite possible that antigen may be destroyed by proteinase formed during storage, and possibly even after the introduction of the filtrate into the tissues during the process of immunization.

Thirdly, it will be seen that anthrax proteinase, or at least that produced by the 'Vollum' strain, unlike streptococcal proteinase is most active in the absence of Na thioglycollate particularly when gelatin is used as substrate.

It would seem therefore that the O/R potential conditions necessary for its activation are the reverse of those required for its production. The necessity for high oxygen tension conditions for the activation of proteinase varies considerably with different filtrates and according to the time of testing. Fresh filtrates in general show no activity in the presence of Na thioglycollate unless grown at a particularly low O/R potential, but after storage considerable activity is evident under reducing conditions. The reasons for this are not clear. It may be that, as it

tested for proteinase activity immediately after filtration and also after storage for 6 and 14 days in a shallow layer at room temperature.

The substrates used were 2·5 per cent powdered milk suspension and 4 per cent gelatin with and without the addition of 0·1 M Na thioglycollate. These figures in each case represent the final concentrations in the substrate filtrate mixtures. The results are given in Table 2.

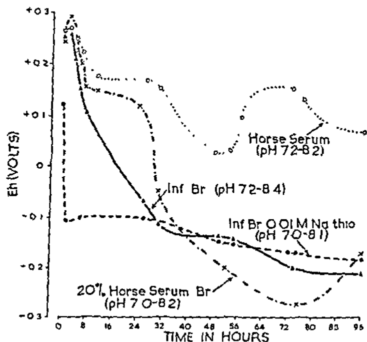


FIG 1. The O/R potentials recorded in static cultures of *B. anthracis* (Vollum) grown in different media

In the first place it will be seen that in these media proteinase production appears to be related to the mean O/R potentials developed during growth (Figure 1). In horse serum, 20 per cent horse serum broth, infusion broth and Na thioglycollate broth, proteinase production increases and the mean O/R potential is lowered in that order. In horse serum, in which the O/R potential is maintained at a relatively high level, little proteinase is formed and then only after prolonged culture. It is

STAPHYLOCOCCAL PROTEINASE AND COAGULASE

It seems possible that *Staph. pyogenes* also may produce an immunizing antigen-destroying enzyme. In the first series of these lectures, Professor Miles (Miles, 1953), discussing various aspects of antibacterial immunity, drew attention to the increasing amount of evidence incriminating coagulase as an important offensive mechanism of staphylococci. He suggested that it only remained to prove the protective value of anticoagulase sera to establish beyond doubt the pathogenic role of coagulase in staphylococcal infections.

Contrary to the findings of earlier workers (Walston, 1935; Smith and Hale, 1944) recent work has shown that coagulase is in fact antigenic (Tager and Hales, 1948; Rammelkamp *et al.*, 1950; Duthie and Lorenz, 1952). Furthermore W. C. Boake from Dr. Gladstone's laboratory, in a paper given before the Pathological Society in January 1953, produced evidence that immunization with coagulase induced the formation of protective antibodies. While Boake's claim that coagulase is an immunizing antigen has not so far as I know been confirmed, it seems appropriate to discuss in this lecture evidence recently obtained that staphylococci produce an enzyme capable of destroying coagulase.

It is well established that coagulase is destroyed by proteolytic enzymes such as pepsin and trypsin (Tager, 1948; Duthie and Lorenz, 1952). Furthermore Lominski and his colleagues (1953) have described a factor believed to be a proteolytic enzyme produced by colonial variants of *Staph. pyogenes* which is also capable of destroying coagulase. During the past two years I have come across evidence of a similar enzyme in the course of an investigation of the action of extracellular enzymes produced by staphylococci on the immunizing activity of the heat-killed vaccines. It appears to be responsible for the rapid disappearance of coagulase from culture filtrates of certain strains on storage, and particularly on incubation. This is demonstrated in the experiment recorded in Table 3, in which filtrates obtained from cultures of various strains, grown for 48 hours at 22° C and used for the preparation of vaccines, were tested for gelatin-liquefying activity and also for their ability to coagulate

ages, proteinase becomes less exacting in its need for a high oxygen tension for activation. Evidence that this may be so is afforded by the fact that some filtrates when fresh produce liquefaction of gelatin only in the upper ⅓ or ¼ of the tube, while on repeating the test liquefaction may be complete. It may also be that proteinase gradually becomes converted into a form requiring the reverse conditions for activation.

On the other hand, it is possible that the anthrax organisms, as appears to be the case with staphylococci, produce two forms of proteinase, one active at a high and the other at a low O/R potential. The development of activity in the presence of thioglycollate on storing the filtrate might be due to the delayed conversion of the latter form from an inactive precursor. It seems probable, however, that this form may be relatively unimportant in so far as destruction of immunizing antigen is concerned, since development is delayed and moreover the O/R potential conditions present in the stored filtrate and after introduction into the tissues are unlikely to favour its activation.

While there is every reason to believe that the immunizing activity of culture filtrates is largely determined by the presence or absence of proteinase, this does not appear to be the only factor concerned. The immunizing antigen appears to be unstable on storage in filtrates apparently containing no proteinase, as in the chemically defined non-protein medium of Wright and his colleagues (Wright, Hedberg and Slein, 1954), though stabilization can be brought about by lyophilization or alum precipitation. It has been suggested that it may be destroyed as the result of oxidation. This does not seem likely in view of Gladstone's observation that it is not destroyed as the result of aerating the culture filtrate (personal communication). The other factor or factors bringing about destruction of the antigen have yet to be determined.

The failure of culture filtrates to immunize is moreover not necessarily due to destruction or inactivation of the antigen. Production may be deficient. As Wright and his colleagues point out, elaboration of the antigen appears to be associated with a particular type of metabolic activity of the organism, to which many media may not be suited.

ascorbic acid, but that this effect is blocked by serum. This may possibly be due to the poising effect of serum on O/R potential. It would appear, therefore, that the activity of coagulase is influenced by O/R potential and is maximal at a relatively high level. The filtrates of fourteen other strains have been tested for proteinase activity in the presence and absence of a reducing agent. Seven gave results similar to that of strain 17, five similar

TABLE 4 The Influence of Na thioglycollate on Proteinase and Coagulase Activity of Culture Filtrates of Different Strains of *Staph. pyogenes*

Strain	Proteinase titre				Coagulase titre			
	Gelatin substrate		Milk substrate		Human		Rabbit	
	+ thio	no thio	+ thio	no thio	+ thio	no thio	+ thio	no thio
17	1/256	1/128	1/512	1/128	—	—	—	—
46	1/8	—	1/32	—	—	—	—	—
B.S.I	—	1/32	—	1/4	—	—	—	—
51	—	—	—	—	—	1/128	—	1/256

— = < 1/2

to strain 46, and two similar to strain 51. The fact that culture filtrates of some strains, even when grown under widely different conditions, consistently show proteolytic activity only in the absence or presence of Na thioglycollate suggests that *Staph. pyogenes* may produce two types of proteinase. One, which for convenience may be referred to as proteinase 'O', like that produced predominantly by the 'Vollum' strain of *B. anthracis* is inactive in the presence of Na thioglycollate. The other, referred to as proteinase 'R', like streptococcal proteinase requires reducing conditions for activation.

The ability of proteolytic filtrates containing one or other or both types to destroy coagulase was studied by mixing with a coagulase-containing filtrate, incubating, and testing the mixtures for loss of coagulase activity. The mixtures were incubated with and without the addition of Na thioglycollate (0 1 M), and controls containing the proteolytic filtrates, previously heated

rabbit plasma immediately after filtration and after incubating for 3 days at 37° C. Apart from the filtrate of strain 46 which contained very little coagulase even when fresh, the rapid disappearance of coagulase from the filtrates on incubation can be seen to be associated with gelatin-liquefying activity.

TABLE 3. The Relationship between *Gelatin-liquefying Activity and* Instability of Coagulase in Culture Filtrates of Different Strains of *Staph. pyogenes*

Strain	Gelatin-liquefying titre	Rabbit coagulase titre of	
		Fresh filtrate	Filtrate after incubation for 3 days, 37°C.
17	$\frac{1}{4}$	$\frac{1}{16}$	—
46	—	$\pm\frac{1}{8}$	—
51	—	$\frac{1}{32}$	$\frac{1}{32}$
139	—	$\frac{1}{32}$	$\frac{1}{16}$
144	$\frac{1}{2}$	$\frac{1}{32}$	—
152	$\frac{1}{4}$	$\frac{1}{8}$	—
221	—	$\frac{1}{32}$	$\frac{1}{32}$
B S.I	$\frac{1}{16}$	$\frac{1}{16}$	—

$- = < \frac{1}{2}$

The proteolytic and coagulase activity of culture filtrates of some of these strains was investigated in greater detail. Cultures were grown for 48 hours at 37° C. and aerated by shaking, in order to obtain as high titres as possible, and the filtrates were tested immediately after filtration (Table 4). Two protein substrates were used, namely milk and gelatin, and the tests were carried out both in the presence and absence of Na thioglycollate (0·2 M). It will be seen that the filtrate of strain 17 is proteolytic both in the presence and absence of the reducing agent, that of strain BS.1 is only active in its absence, and that of strain 46 only in its presence. Essentially the same results were obtained with both substrates. The filtrate of strain 51 showed no proteinase activity and is the only one to coagulate human or rabbit plasma. It will be noted furthermore that coagulase activity is inhibited by Na thioglycollate. Tager (1948) has recorded that coagulase is also exceedingly sensitive to

ascorbic acid, but that this effect is blocked by serum. This may possibly be due to the poising effect of serum on O/R potential. It would appear, therefore, that the activity of coagulase is influenced by O/R potential and is maximal at a relatively high level. The filtrates of fourteen other strains have been tested for proteinase activity in the presence and absence of a reducing agent. Seven gave results similar to that of strain 17, five similar

TABLE 4 The Influence of Na thioglycollate on Proteinase and Coagulase Activity of Culture Filtrates of Different Strains of *Staph. pyogenes*

Strain	Proteinase titre				Coagulase titre			
	Gelatin substrate		Milk substrate		Human		Rabbit	
	+ thio	no thio	+ thio	no thio	+ thio	no thio	+ thio	no thio
17	$\frac{1}{256}$	$\frac{1}{128}$	$\frac{1}{512}$	$\frac{1}{128}$	—	—	—	—
46	$\frac{1}{8}$	—	$\frac{1}{32}$	—	—	—	—	—
B.S I	—	$\frac{1}{32}$	—	$\frac{1}{4}$	—	—	—	—
51	—	—	—	—	—	$\frac{1}{128}$	—	$\frac{1}{256}$

– = < ½

to strain 46, and two similar to strain 51. The fact that culture filtrates of some strains, even when grown under widely different conditions, consistently show proteolytic activity only in the absence or presence of Na thioglycollate suggests that *Staph. pyogenes* may produce two types of proteinase. One, which for convenience may be referred to as proteinase 'O', like that produced predominantly by the 'Vollum' strain of *B. anthracis* is inactive in the presence of Na thioglycollate. The other, referred to as proteinase 'R', like streptococcal proteinase requires reducing conditions for activation.

The ability of proteolytic filtrates containing one or other or both types to destroy coagulase was studied by mixing with a coagulase-containing filtrate, incubating, and testing the mixtures for loss of coagulase activity. The mixtures were incubated with and without the addition of Na thioglycollate (0·1 M), and controls containing the proteolytic filtrates, previously heated

to 80° C. for one hour to inactivate the enzyme, were also set up. The results are recorded in Table 5. It should perhaps be mentioned that the amount of Na thioglycollate present in some of the mixtures after incubation was insufficient to affect the titration of coagulase apart from producing a slight prozone of inhibition in one or two cases.

TABLE 5. The Destructive Action of Proteolytic Filtrates on Coagulase Activity of Strain 51 filtrate

Culture filtrate of strain 51 incubated with filtrates below (grown 48 hrs. 37°C., shaken)	Proteinase titres of mixtures before incubation		Coagulase titres of mixtures after incubation	
	+ thio ('R')	– thio ('O')	*Human	†Rabbit
17. – thio (2 days 37°C.)	1/128	1/64	—	—
17. + thio (2 days 37°C.)	1/256	1/64	1/16	1/16
B.S.I. – thio unheated (10 days 37°C.)	—	1/8	—	1/4
B.S.I. + thio unheated (10 days 37°C.)	—	—	1/64	1/128
B.S.I. – thio heated (10 days 37°C.)	—	—	1/64	1/256
46. – thio unheated (10 days 37°C.)	1/16	—	1/64	1/256
46. + thio unheated (10 days 37°C.)	1/16	—	1/64	1/256
46. + thio heated (10 days 37°C.)	—	—	1/64	1/128
139. + thio unheated	1/16	—	1/64	1/256
Fresh broth – thio	—	—	1/128	1/256
Fresh broth + thio	—	—	1/64	1/256

Before incubation * 1/256, † 1/512
– = < 1/4

It will be seen that while the unheated proteinase 'O' containing filtrates produced a marked reduction in coagulase titre, those containing the 'R' form were inactive, apart from that of strain 17 which contained both types. A similar result was obtained on heating the mixtures at 80° C. for 1 hour before testing, in order to destroy the proteinase present, while leaving the activity of the heat-resistant coagulase unimpaired.

There appears, therefore, to be a marked difference in the ability of these two presumably distinct forms of proteinase to attack coagulase. The reasons for this have yet to be determined.

The Search for Antigen-destroying Enzymes in other Bacterial Species

It seems most improbable that the production of antigen-destroying enzymes in cultures of bacteria is confined to those species in which it has so far been demonstrated. Nor is there any reason to assume that only protein antigens are liable to be destroyed in this way. In this connection I might mention the destruction of hyaluronic acid present in the capsular material produced by some strains of haemolytic streptococci as the result of the action of autogenous hyaluronidase (Seastone, 1943; McLean, 1941, 1942). This substance, however, though possibly an important offensive factor is non-antigenic. There is, furthermore, no reason to assume that antigens may only be destroyed by enzymes liberated from the cells during growth. It seems likely that the process may occur much more frequently than is generally realized and may be responsible in a number of instances for the irregular or inconstant production of immunizing antigens in bacterial cultures and their instability in vaccine or culture filtrate.

For example, some of the peculiar characteristics of the Vi antigen of the typhoid bacillus may be explained on the assumption that it is subject to destruction by enzyme action. It is well known that this antigen is present most abundantly in relatively young cultures of *Salm. typhi*, and that it tends to disappear in the later stages of growth, i.e. after about 10–12 hours. It is also known that cultures of some strains, such as Ty.2, retain it for a longer period than others, and in cultures of other Salmonella species which produce it, such as *Salm. ballerup*, it may be quite stable. This behaviour may well be due to the fact that these strains produce varying amounts of a Vi-antigen-destroying enzyme.

Furthermore, Felix and his colleague (Felix and Petrie, 1938; Felix, 1941) showed that while the immunogenicity of the T Vi antigen was lost following heat treatment at 60° C. for one

hour, it was largely unimpaired in alcohol-killed suspensions. Later (Felix, 1952) it was shown that the degree to which it was lost following heat treatment varied with the strain, and as before in some strains, particularly those of other Salmonella species, it was relatively stable. According to Peluffo (Peluffo, 1941), the normally heat-sensitive T Vi antigen is highly resistant to heat when the bacilli are suspended in absolute alcohol or acetone or when they are dehydrated *in vacuo*. In explanation of these observations I would suggest that the T Vi antigen may be destroyed not so much by the heat treatment as by enzyme action which continues after heating and may be accelerated by it, and that this process is partially if not completely arrested by alcohol. Felix, in a recent paper (1952), concludes that 'the physicochemical behaviour of . . . the T Vi antigen may vary as the result of the presence in, or absence from, the bacterial cell of some other substance which itself may be either antigenic or non-antigenic'. If such a Vi-antigen-destroying enzyme does exist and is responsible for the effects I have described, it may well remain bound to the cells and have to be extracted in some way before it can be studied.

Conclusions

But to return from surmise to the more definite conclusions to be drawn from the experiments I have described. Besides affording evidence of the destruction of certain antigens by bacterial enzymes, they have also shown that factors such as O/R potential are of considerable importance in regulating the activity of these enzymes. The antigenic make-up of an organism growing under the relatively low O/R potentials developed in bacterial cultures may therefore be quite unlike that of the same organism growing *in vivo*, where the oxygen tension conditions are likely to be very different. This may possibly be one reason for the superior results sometimes obtained with living, as compared with dead vaccines.

In conclusion it may be useful if I give a summary of the methods used in these investigations, including those that may prove successful in the study of possible intracellular antigen-destroying enzymes.

The first step is to attempt to produce a highly immunizing antigen either in the vaccine, as with *Str. pyogenes*, or in culture filtrate, as with *B. anthracis*, by the usual methods, given in Table 6. The next step is to reverse the procedure and, by changing the strain or by slight modifications in growth conditions, to obtain, if possible, preparations which are completely non-immunizing. If successful, the next step is to determine whether this lack of immunizing activity is in fact due to

TABLE 6. Stages in the Investigation of Immunizing Antigen-destroying Enzymes (I.A D.E.) of Bacteria

I. Production of *fully immunizing* vaccine (or culture filtrate) by varying
   (*a*) strain (including artificially induced variants)
   (*b*) medium
   (*c*) growth conditions (e.g. Eh and pH)
   (*d*) method of killing vaccine
   (*e*) conditions of storage.

II Production of *non-immunizing* vaccine (or culture filtrate) by modifying (*a*), (*b*) or (*c*) above.
   Preparation of filtrate (or cell extract) from *non*-immunizing vaccine culture.

III. Treatment of *immunizing* vaccine (or culture filtrate) from I with culture filtrate (or cell extract) from II by incubating (at varying Eh and pH), and tests of treated product for destruction of I.A by
   (*a*) immunization experiments
   (*b*) serological methods.

IV. Parallel tests of filtrates or extracts from II with substrates of same chemical nature as immunizing antigen (where known).

V. Establishment of correlation between presence of I A and absence of I.A.D E in a culture (and vice versa).

VI. Confirmatory tests with purified I.A D.E

the production of an immunizing antigen-destroying enzyme. This is done by treating the immunizing antigen with the culture filtrate from the non-immunizing preparation, including also heated filtrates as controls, and testing the treated product for destruction of the antigen by active immunization, serological or other tests. If negative results are obtained, it may be due to the fact that the enzyme is only developed intracellularly. Preparations should therefore be made of intracellular products by extracting, lysing or disintegrating the cell, and these should be tested in the same way.

When the chemical nature of the immunizing antigen is

hour, it was largely unimpaired in alcohol-killed suspensions. Later (Felix, 1952) it was shown that the degree to which it was lost following heat treatment varied with the strain, and as before in some strains, particularly those of other Salmonella species, it was relatively stable. According to Peluffo (Peluffo, 1941), the normally heat-sensitive T Vi antigen is highly resistant to heat when the bacilli are suspended in absolute alcohol or acetone or when they are dehydrated *in vacuo*. In explanation of these observations I would suggest that the T Vi antigen may be destroyed not so much by the heat treatment as by enzyme action which continues after heating and may be accelerated by it, and that this process is partially if not completely arrested by alcohol. Felix, in a recent paper (1952), concludes that 'the physicochemical behaviour of . . . the T Vi antigen may vary as the result of the presence in, or absence from, the bacterial cell of some other substance which itself may be either antigenic or non-antigenic'. If such a Vi-antigen-destroying enzyme does exist and is responsible for the effects I have described, it may well remain bound to the cells and have to be extracted in some way before it can be studied.

Conclusions

But to return from surmise to the more definite conclusions to be drawn from the experiments I have described. Besides affording evidence of the destruction of certain antigens by bacterial enzymes, they have also shown that factors such as O/R potential are of considerable importance in regulating the activity of these enzymes. The antigenic make-up of an organism growing under the relatively low O/R potentials developed in bacterial cultures may therefore be quite unlike that of the same organism growing *in vivo*, where the oxygen tension conditions are likely to be very different. This may possibly be one reason for the superior results sometimes obtained with living, as compared with dead vaccines.

In conclusion it may be useful if I give a summary of the methods used in these investigations, including those that may prove successful in the study of possible intracellular antigen-destroying enzymes.

GLADSTONE, G. P. (1948). *Brit. J. exp. Path.* **29**, 379.

GRABER, P. and STAUB, A. M. (1946). *Ann. Inst. Pasteur*, **72**, 534.

HEWITT, L. F. (1950). 'Oxidation Reduction Potentials' in *Bacteriology and Bio-chemistry*, 6th ed. E. & S. Livingstone, Edinburgh.

[illegible]

[illegible] 53, p. 192.

MILLER, W. S., CLARK, D. L. and DIERKHISING, O. C. (1951). *Amer. J. trop. Med.* **31**, 535.

OAKLEY, C. L., WARRACK, G. H. and VAN HEYNINGEN, W. E. (1946). *J. Path. Bact.* **58**, 229.

PELUFFO, C. A. (1941). *Proc. Soc. exp. Biol., N.Y.* **48**, 340.

RAMMELKAMP, C. H., HEZEBICKS, M. M. and DINGLE, J. H. (1950). *J. exp Med.* **91**, 295.

SCHILLING, S. J. (1927) *J. Amer. Vet. med. Ass.* **72**, 300.

SEASTONE, C. V. (1943). *J. exp. Med.* **77**, 21.

SMITH, W. and HALE, J. H. (1944). *Brit. J. exp. Path.* **25**, 101.

SOBERNHEIM, G. (1913). *Handbuch der pathogenen Mikroorganismen* Edit. W. Kolle and A. von Wassermann (2nd. edn., Fischer, Jena), **3**, 583.

STAMP, LORD (1953). *Brit. J. exp. Path.* **34**, 347.

STAMP, T. C. and HENDRY, E. B. (1937). *Lancet*, **1**, 257.

STRANGE, R. E. and BELTON, F. C. (1954). *Brit. J. exp. Path.* **35**, 153.

TAGER, M. (1948). *Yale. J. Biol. Med.* **20**, 487.

TAGER, M. and HALES, H. (1948). *J. Immunol.* **60**, 475

TILLETT, W. S. and GARNER, R. L. (1933). *J. exp. Med.* **58**, 485.

TODD, E. W. (1947). *J. exp. Med.* **85**, 591.

WALSTON, H. D. (1935). *J. Hyg.* **35**, 549

WHITE, P. B. (1946). *Brit. J. exp. Path.* **27**, 356.

WRIGHT, G. C., HEDBERG, M. A. and FEINBERG, R. J. (1951). *J. exp. Med.* **93**, 523

WRIGHT, G. C., HEDBERG, M. A. and SLEIN, J. B. (1954). *J. Immunol.* **72**, 263

known the enzyme concerned may also be studied by additional tests, using more readily available alternative substrates.

Again, as has been demonstrated in the studies of staphylococcal coagulase and proteinase, the establishment of a correlation between presence of enzyme and absence of antigen and vice versa in a culture affords additional evidence of antigen-destroying activity, and may give some indication as to whether the action of the enzyme is solely responsible for lack of immunizing activity. Lastly, it is necessary to isolate the enzyme in as chemically pure a form as possible and to carry out confirmatory tests with the purified product. But this is likely to be a long and laborious task, as was found to be the case with streptococcal proteinase.

The study of antigen-destroying enzymes is obviously only in its infancy, and much more experimental work is required before it is possible to assess their importance as factors influencing the immunizing activity of vaccines in general.

I believe, however, that the evidence is sufficiently promising to warrant further extensive research along these lines.

REFERENCES

Bail, O. (1904). *Zbl Bakt., Abt.* 1, 36, 266.
[illegible] 144.
[illegible] *ubjects*, Edit.
[illegible] York, 2, 49
[illegible]
[illegible] R. J. (1947) *J. infect. Dis* 80, 14.
Duthie, E S and Lorenz, L. (1952). *J. gen Microbiol* 6, 95.
Elliott, S. D (1945). *J. exp. Med.* 81, 573
Elliott, S. D (1950). *J. exp Med.* 92, 201.
Elliott, S. D. and Dole, V. P (1947) *J exp. Med.* 85, 305.
Felix, A. (1941). *Brit med J* 1, 391.
Felix, A. (1952). *J. Hyg.* 50, 515.
Felix, A. and Petrie, G. F. (1938). *J Hyg.* 38, 673.
Frobisher, M. Jr. (1926). *J exp. Med.* 44, 777.
Gladstone, G P (1946). *Brit. J exp Path* 27, 394.

chick embryo for viruses and in ordinary culture media for fungi.

An antibiotic will naturally be useless in a condition which is not infective at all, and it seems appropriate to point out that fever is not invariably the result of bacterial action, since this symptom appears sometimes to be considered an indication in itself for antibiotic treatment. Some conditions have the features of an infection, and yet the part played in them by bacteria is only secondary: among these I am inclined myself to place ulcerative colitis, in which on the whole antibiotic treatment has been disappointing. How secondary a part, if indeed any, bacteria may play in an ulcerative process was brought home forcibly to me by two patients recently in St. Bartholomew's Hospital under the care of my colleague, Dr. E. F. Scowen. One of these was a boy with extensive ulceration of the mouth and fauces, in which no specific bacteria could be found, and the second a man diagnosed as having ulcerative colitis but giving a history of a similar mouth condition at the onset: the condition in both was quite irresponsive to antibiotics, but cleared up rapidly and completely when cortisone was given.

It should also be remembered that although a condition may be clearly infective in nature, its existence may depend on an underlying structural defect: unless this can be remedied permanent cure cannot be hoped for. Almost all authors who have studied the chemotherapy of urinary infections have emphasized that in the presence of calculi or any form of obstruction relapse is certain to occur even though the urine may for the time be sterilized. Another example of the effect of stagnation in predisposing to infection is bronchiectasis: chemotherapy can afford complete temporary relief of symptoms, but—except perhaps in the mildest cases in children (Franklin and Garrod, 1953)—restoration of the *status quo ante* occurs eventually. Many examples could be cited of conditions in which antibiotic treatment is only an adjunct to surgery, and cannot replace it. Closed foci of suppuration still require drainage and large areas of tuberculous caseation or any tissue which has undergone necrosis must be removed before resolution can take place.

XVII

Causes of Failure in Antibiotic Therapy

LAWRENCE P. GARROD

I AM not very sure that the subject of this lecture qualifies for inclusion in a series with so high-sounding a title as 'The Scientific Basis of Medicine'. It involves a somewhat prosaic story of mistakes, most of which can easily be avoided, although *admittedly the avoidance of some of them calls for rather* deeper thinking and a knowledge of bacterial behaviour which is at least not universal.

The causes of failure in chemotherapy may obviously be found in the patient, in the choice of the drug and/or the mode of its use, or in the micro-organism itself. I propose to consider them in that order.

Conditions Insusceptible to Antibiotic Treatment

Most pathogenic organisms are sensitive to one or more of the major antibiotics, but there are still important exceptions. It is interesting that these are found at the extremes of the scale of microbic dimensions. At one end are the smaller viruses; so far, in spite of immense efforts and some interesting findings, there is no prospect that any antibiotic will influence the course of most virus diseases. At the other end are the fungi, and with the important exception of actinomycosis, none of the systematic mycoses is amenable to chemotherapy, unless the recently discovered antibiotic 'nystatin' should prove this to be untrue. So resistant are both these classes of organism to the major antibiotics that the latter can be used in high concentration to facilitate their isolation from contaminated material, in the

demands treatment with one of the newer antibiotics: the nature of the infection can usually be determined by simply staining a film of the sputum, and there need therefore be no delay.

Two other parts of the body in which infection may be due to any of a number of bacteria are the meninges and the urinary tract. There is no condition in which expert laboratory help is more urgently necessary than meningitis. Both here and in the urinary tract the infection may be caused by a species of Gram-negative bacillus, such as *Ps. pyocyanea*, which is highly resistant to most chemotherapeutic agents. The antibiotics to be considered in such infections should include polymyxin: *Ps. pyocyanea* and other coliform bacilli are highly sensitive to it, and in the doses required it is non-toxic as well as effective (Jawetz, 1952).

Sometimes no single drug will succeed, and a combination of two is indicated. This is much less often necessary than the present popularity of 'combined' treatment would suggest, but it may be life-saving when the organism is comparatively resistant and must be completely eradicated, as in *Str. faecalis* endocarditis. Gunnison and Jawetz (1950) have shown that penicillin even in optimal concentration will not exterminate this organism *in vitro*, but if streptomycin be added, even in a concentration which acting alone permits growth, the culture will be sterilized. This combination of drugs is in fact the only successful treatment for this form of endocarditis (Robbins and Tompsett, 1951; Cates, Christie and Garrod, 1951). The sensitivities to antibiotics of four strains of *Str. faecalis* from cases of endocarditis treated at St. Bartholomew's Hospital are stated in Table 2. All the patients responded to this treatment (although not all recovered) and several of them had had previous prolonged but ineffective treatment with the antibiotics —aureomycin or in one case chloramphenicol—to which their streptococci were most sensitive by the ordinary method of test (Garrod, 1953).

Penicillin and strept [illegible]

been proposed by Jawetz and Gunnison (1953); suffice it to say

WRONG CHOICE OF ANTIBIOTIC

The bacteriological diagnosis, as in, for instance, carbuncle or gonorrhoea, is sometimes as self-evident as the clinical, but in many forms of disease it can only be ascertained with certainty in the laboratory, and if it is simply guessed at, and the guess proves to be wrong, the antibiotic chosen may be ineffective. Pneumonia, for example, may be caused by any of five species of bacteria, apart from viruses. If penicillin be the drug chosen, it will be seen from Table 1 that it is likely to be effective in two

TABLE 1. Action of Penicillin on Bacteria Causing Pneumonia

Species	Minimum inhibitory concentration (units/ml.)
Str. pneumoniae	0 01
Str. pyogenes	0 01
Staph. pyogenes[1]	0 02
H. influenzae	0·5–5 0
Kl. pneumoniae	5–100

[1] Many strains much more resistant

of these infections, including of course the commonest (pneumococcal), and to be ineffective in infections due to the two Gram-negative bacilli, while in a staphylococcal infection it may or may not be according to whether the strain is sensitive or resistant. It has been claimed for the broad spectrum antibiotics, and in particular aureomycin (Herrell, 1949) that all kinds of pneumonia, including those due to viruses, can be successfully treated with them, and it was thus to be expected that routine use would yield better results than treatment with penicillin. The Medical Research Council Trial (Report, 1951) of aureomycin and chloramphenicol in pneumonia did not bear this out: these drugs gave no better over-all results than previously standard treatment (usually penicillin) and were disliked by the patients because of their effects on the digestive tract. This series included only 3 *Kl. pneumoniae* and 1 *H. influenzae* infection, and all of these happened to be in the groups treated with aureomycin or chloramphenicol. A pneumonia due to either of these organisms, particularly the former, urgently

neglect of this preliminary treatment is a common cause of failure, and I have seen an illustration of this quite recently.

Defects in Dosage

There are several reasons why inadequate dosage of the newer antibiotics is uncommon. The obvious one is that oral administration presents no difficulties: others are the well-sustained blood levels after infrequent doses and the fact that the smaller the dose the greater the proportion of it absorbed.

Penicillin presents a different problem. Unless it be admitted that oral administration is satisfactory—and about this there is some doubt in more serious infections and those less readily susceptible to treatment—the difficulty is that of providing for repeated injections. It is usually feasible in any circumstances to give these once daily, and some years ago it was frequently said and generally supposed that a sufficiently large dose of sodium penicillin given once daily was adequate. A colleague and I (Garrod and Shooter, 1950) gave 500,000 units of sodium penicillin to each of 32 subjects and assayed the drug in the blood by an exceedingly sensitive and accurate method at intervals after the injection. The average period during which penicillin could be detected was slightly less than 8 hours. It has been shown by Eagle and Musselman (1949) that a period of up to 4 hours elapses after non-lethal exposure to penicillin before bacterial growth is resumed. It therefore seems reasonable to suggest that this dose should be given twice daily and not once for full therapeutic effect.

It is often argued, on the other hand, that the high initial blood concentration produced with sodium penicillin favours diffusion into foci of infection, and that the drug persists in such foci long after it has disappeared from the blood. The first part of this argument is probably true, the second true only of foci where there is a collection of exudate such as an abscess. In an infected area where the circulation is normal or increased and no tissue breakdown has occurred, there is no reason why penicillin should accumulate and persist, and the ingenious experiments of Eagle, Fleischman and Levy (1953a), who assayed the penicillin content of infected muscle at intervals

that antagonism may result from the combination of a bacteriostatic drug with a bactericidal one, the latter only exerting this effect on multiplying organisms. There is plentiful evidence of such an interfering effect (e.g. by chloramphenicol or aureomycin, which are 'static, with penicillin, which is 'cidal) both *in vitro* and in experimental infections in mice. There is also one

TABLE 2. Antibiotic Sensitivities of 4 Strains of *Str. faecalis* Isolated from Cases of Endocarditis, St. Bartholomew's Hospital, 1950–2

		Minimum Inhibitory Concentrations (μg./ml.)			
Sex	Age	Penicillin	Streptomycin	Aureomycin	Chloramphenicol
F	36	2·5	30	0·15	20
M	57	1·8	30	0·1	6
M	63	2 5	125	0 15	2·5
F	25	2·5	10	0·3	0 6

clear example, in special conditions, in the clinical field. Lepper and Dowling (1951) obtained a recovery rate of 70 per cent in pneumococcal meningitis by treatment with penicillin alone, in intramuscular doses of 2,000,000 units two-hourly. When aureomycin was added to this treatment recoveries fell to 21 per cent. Both drugs evidently reached the cerebrospinal fluid in the critical rather low concentrations necessary for antagonism, and the meninges are a situation in which a full bactericidal effect is necessary.

DEFECTS IN ADJUVANT TREATMENT

A minor cause of failure is in not providing the conditions under which an antibiotic can exert its full effect. An example of this, apparently not even yet generally familiar, is failure to alkalinize the urine before giving streptomycin for a urinary tract infection. This antibiotic depends more than any other antibiotic for its effect on pH, and alkalinity favours it. The pH almost everywhere in the body (collections of pus are one exception) is on the alkaline side except in the urine. It is therefore imperative that the urine should be rendered constantly alkaline, as judged by examining the first morning specimen, by giving full doses of sodium bicarbonate, before treatment is begun. The

system of administering penicillin would have to be reconsidered.

Penicillin treatment may fail, wholly or in part, because doses are inadequate or intervals between them too long, but also because treatment is stopped too soon. If the course given for streptococcal angina, including scarlet fever, is too short, recurrence is common. Diseases for which exceptionally long courses are imperative are subacute bacterial endocarditis and actinomycosis.

Misguided Prophylaxis

There is much difference of opinion about the indications for antibiotic prophylaxis, including a strong body of well-informed opinion that this proceeding is resorted to much too freely. When it is, there should be a clear idea of its aim. Few would deny that penicillin should be used as cover for dental extraction in a rheumatic subject, but is the object of this to destroy bacteria entering the circulation, or to destroy them beforehand in the mouth itself? If the former, the first dose should be given immediately before extraction: if the latter, presumably several days beforehand. In fact the former should be the object: sterilization of the mouth is impossible, and attempts to do so merely eliminate the sensitive part of the flora and so encourage the growth of resistant organisms. The grave danger of thus selecting as a cause of subsequent infection an organism insusceptible to the preferred antibiotic is well illustrated by the history of a recent patient at St. Bartholomew's Hospital (Cates, Christie and Garrod, 1951) A man who had been cured of subacute bacterial endocarditis with penicillin five years earlier had to have ten teeth removed. He was admitted to hospital and given 500,000 units of penicillin 4 times a day for 6 days, extractions being done 2 and 4 days after the beginning of this course. He promptly developed a second attack of bacterial endocarditis due to a streptococcus 200 times more resistant to penicillin than that causing his original infection, and is fortunate to be still alive, although with damage to the vestibular branch of his eighth nerve, after a prolonged course of heavy doses of penicillin and streptomycin.

after a single therapeutic dose, prove that it does not. Many of the conditions for which penicillin is used, such as cellulitis, pneumonia, acute catarrh, are of this general nature—acute inflammation with hyperaemia and no tissue breakdown.

This vexed question of how often sodium penicillin should be given may well be regarded as having been settled by an elaborate and well-designed series of experiments reported by Eagle, Fleischman and Levy (1953b). Tables 3 and 4 are

TABLE 3. Streptococcal Infection in Mice Treated with 4 Doses of Penicillin
(*Data from Eagle et al.*, 1953b)

Interval between injections (hours)	$\frac{3}{4}$	$1\frac{1}{2}$	3	6	12	24
Total dose (mg /kg.) required to cure	68	31	3	6	7 6	210

extracted from the large mass of data in this paper, and both show in different ways that there is an optimum interval for a given dose: if this is either shortened or exceeded, not only must a greater amount of the drug be given to achieve the same effect, but—and this applies to a lengthened interval only—the duration of treatment necessary for cure is greatly extended.

TABLE 4. *Streptococcal Infection in Mice Treated with Penicillin in Doses of 3 2 mg./kg. (= ± 250,000 units in man)*
(Data from Eagle *et al.*, 1953b)

Interval between dose (hours)	$\frac{3}{4}$	$1\frac{1}{2}$	3	6	12	24
Treatment time necessary for cure	*$4\frac{1}{2}$–6*	*6–9*	*6–9*	*12–18*	*36–48*	*>192*
I.e., number of doses	6–8	6–8	3–4	2–3	3–4	>8

It is much easier to secure a continuous effect by using procaine penicillin, and this form of the drug has been increasingly popular, particularly for domiciliary and out-patient treatment. In some quarters its use is now being discouraged because it is said to be more liable to produce shock, sometimes fatal, in sensitized patients. If the severe restrictions on its use proposed by Kern and Wimberley (1953) were to be adopted, the whole

some extent in the entire population of large areas. This general change chiefly occurs and has been extensively studied in Gram-negative, especially coliform, bacilli, and in staphylococci, and it seems possible that these two varieties of organism may be the chief ultimate survivors in the contest now in progress between pathogenic bacteria and a succession of antibiotics. That coliform bacilli are more frequently resistant to streptomycin than in the past is common experience, and the rapid increase in

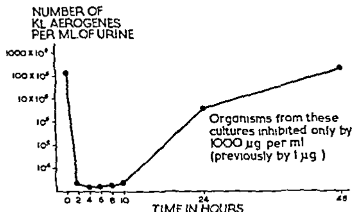


FIG. 1. Effect of streptomycin (0.5 gm four-hourly starting at 0 hours) on bacterial content of urine in a case in which treatment failed.

prevalence of resistant strains in a country where the drug has been widely used is illustrated in the findings of Romansky and his colleagues (1951). The same change is following the extensive use of aureomycin, chloramphenicol and terramycin (Thomson, 1952).

The whole of this heterogeneous group of bacteria is perhaps of less importance as a cause of antibiotic-resistant infections than one other single species, *Staph. pyogenes*. This organism is not only very frequently penicillin-resistant, but in the United States is now commonly resistant to the newer antibiotics as well. The observations of Needham and Nichols (1953) and Kirby and Ahern (1953) make it clear that the proportion of resistant strains in any given population depends directly on the extent to which the drug has been used in the immediate past.

ACQUIRED BACTERIAL RESISTANCE

The last cause of therapeutic failure with which I propose to deal is perhaps the first which would occur to many people's minds, and it is certainly growing steadily in prominence. Acquired bacterial resistance to the antibiotic used, whether of moderate or of so high a degree as to render the drug useless, has to be considered from two points of view, that of the patient in whom the change occurs, and that of the community in which these resistant strains become prevalent.

In the individual patient, the use of penicillin involves little risk of this kind. Some important bacteria are incapable of developing penicillin-resistance in any circumstances: others can do so *in vitro* but in fact do not *in vivo*. I have never seen the organism causing an infection acquire resistance during its treatment with penicillin except in a few cases of actinomycosis and of subacute bacterial endocarditis: here the increases were moderate and occurred during prolonged treatment. Penicillin resistance in staphylococci is of course a serious problem, but its prevalence is due not to habituation but to selection, or in the individual to substitution. The position is very different with streptomycin: all bacteria can acquire a high degree of resistance to this drug with astonishing rapidity. In Figure 1 is shown the course of events in a case of urinary tract infection unsuccessfully treated with streptomycin. The bacterial count in the urine fell within two hours of the first dose to a low level, but persisted there and rose again overnight, the organism cultivated on the following morning being already 1,000 times more resistant than it had been 24 hours earlier. That the same change can occur, although more slowly, in tubercle bacilli is well known, and the combinations of other drugs with streptomycin now universally employed in treating tuberculosis are designed so far as possible to prevent it. Resistance to the newer antibiotics can also be acquired by most [illegible] isms, but the process is much slower [illegible] f difficulty in the [illegible]

[illegible] and strains which possess or [illegible] it pass from person to person, eventually attaining widespread prevalence, particularly in hospitals, but to

this process had been accepted as the method of human infection. But, as soon as experimental inoculations for the malarial therapy of G.P.I. were carried out at the Malaria Reference Laboratories at Horton, it became evident that blood-inoculated malaria (trophozoites) had a different natural history from mosquito-transmitted infection (sporozoites). Though parasites could be found in the blood of the recipient of trophozoites from the moment of infection, and fever developed quite soon, sporozoites on the other hand disappeared after an hour and it was a week or nine days before the clinical attack developed. It was evident that sporozoites, injected either by the mosquito or the hypodermic needle, underwent an exo-erythrocytic phase before they could cause parasitaemia. The original work with *P. cynomolgi* of monkeys was confirmed for *P. vivax* and *P. falciparum* in man, and has recently been demonstrated for *P. ovale* also Infection starts with an exo-erythrocytic cycle in the liver So it was that in 1948, the London School of Hygiene and Tropical Medicine converted Ross's bicycle into Shortt's tricycle, and placed the life cycle of the malaria parasite on a more stable basis. It was not long before the same team of tricyclists were able to add a fourth wheel to account for the relapsing phase of *P. vivax* and presumably of *P. malariae* and *P. ovale*. For, unlike *P. falciparum*, these parasites undergo further stages of exo-erythrocytic development in the liver associated with the return of parasitaemia and the clinical relapse.

We are thus confronted by a four-phase mechanism which can be conveniently likened to a car on four wheels. The near front wheel represents the process of asexual schizogony in the blood stream, with its cycle of 48 to 72 hours. The near side rear wheel symbolizes sexual sporogony in the mosquito's stomach, the speed of which depends on the atmospheric temperature which gears up or down the rate of arrival of sporozoites in the salivary glands.

Malaria transmission occurs via the rear axle to the off-side

XVIII

Antimalarial Drugs

CLEMENT C. CHESTERMAN

The Different Stages of the Malaria Parasite

In order to get an historical and parasitological background let us review the story of the discovery of the life history of the malaria parasite.

From the millennia of myth and mystery, the first solid fact in the pathology of malaria was the discovery in 1846 by Meckel, of little brown-black dots of pigment in the viscera of patients with malaria. In 1872, Delafield demonstrated that this pigment was surrounded by transparent, granular protoplasm. Laveran, in 1880, identified these bodies in the red blood-cells of malaria patients, and Marchiafava in 1884, using the oil immersion lens, noted that they were amoeboid.

The malarial cycle of asexual erythrocytic schizogony was established and one could watch the wheel turn full circle in 48 or 72 hours according to the periodicity of the parasite.

Ten years later, Manson, watching the exflagellation of the male gametocytes in shed blood, suggested that it was the mosquito which provided the milieu for the evolution of this mechanism outside man. 'Look for the parasite's dung in the mosquito's stomach' was his advice to Ross, referring to the telltale pigment, and, sure enough, Ross found it in the stomach of his dapple-winged mosquito. Manson imagined that the mosquito thus infected might conveniently die in water and if this were drunk by human beings, transmission might occur.

But this penny-farthing hypothesis was soon converted by Ross into a modern bicycle when he found the sporozoites in the salivary glands and infected his sparrows from a mosquito

3. *The Primary Clinical Attack.* The pre-erythrocytic merozoite from the liver invades the red cell and grows through trophozoite schizont and merozoite stages, each 48 or 72 hours long, and this asexual schizogony is continued till immunity develops. A drug acting at this stage is called a schizonticide or suppressive drug in that it abolishes the clinical attack. In this class are quinine and the other cinchona alkaloids, mepacrine, chloroquine and the other 4-aminoquinolines. Proguanil and pyrimethamine only act slowly at this stage, and the 8-aminoquinolines hardly at all in safe doses, with the possible exception of primaquine.

4. *The Latent Phase* follows the primary clinical attack and is asymptomatic. During it gametocytes appear in the blood and a drug which causes them to degenerate and disappear is called gametocidal, or more strictly gametocytocidal. Here again the 8-aminoquinolines, reinforced with quinine or chloroquine, are the most effective. Obviously all schizonticides have an effect on gametocyte control by preventing their formation.

5. *The Relapse.* During the latent period the process of asexual schizogony is reduced to sub-clinical proportions. In *P. falciparum* malaria there may be a recrudescence of parasitaemia due to a revival of this process. But in the other three forms of malaria latency is regularly interrupted by relapses due to the emergence of persistent tissue forms from the liver. A *Curative drug* is one which will prevent this happening, and again the 8-aminoquinolines, primaquine *par excellence*, combined with quinine or chloroquine, are effective. Other drugs which help to prevent the relapse do so by scotching the exo-erythrocytic merozoites when they reappear in the blood, e.g. mepacrine.

6. *Transmission.* This depends on sporogony in the mosquito which embraces the processes of exflagellation, fertilization and oocyst maturation. This latter fails if sufficient, albeit minute quantities of certain drugs are ingested by the mosquito along with its blood meal. Such a drug can be called sporonticidal and proguanil and pyrimethamine taken regularly by man, the intermediate host, prevent maturation of oocysts in the definitive host, the mosquito, and bring transmission to a standstill.

rear wheel where sporozoites injected by the bite start a pre-erythrocytic cycle in the liver cells. A re-invasion of the liver by the resulting schizonts in the relapsing malarias adds a fourth wheel.

The motive force of the malarial chariot is a blended mixture of the mating instinct of the male mosquito and the blood lust of the female. If these are cooled down to below 65[1] F., transmission comes to a standstill; neither of the back wheels can revolve.

This man–mosquito–man mechanism has menaced the highways and byways of human progress. It is interesting to speculate how it started up. Did it get a first push off from some grinning chimpanzee who afterwards hopped up behind and curled up like a spare wheel? It seems certain that *P. malariae* at least is common both to man and the chimpanzee and it would not be surprising if gorillas were found to harbour other species in common with man. The fear of an animal reservoir need not worry us much for we are rarely fellow travellers.

Our concern is to see how drugs can wreck this car or jam its mechanism.

The Natural History of Malarial Infection

This can be divided into the following stages:

1. *Infection,* during which sporozoites are injected into the blood stream. It is probable that only those which reach the liver survive. A drug which would kill sporozoites during their brief stay in the circulation would be a *True Causal Prophylactic*. Unfortunately none is known

2. *Prepatent* [illegible]
the ons[illegible] the clinical attack. It lasts [illegible] days for *P. falciparum* and nine for *P. vivax* and *P. ovale*. This corresponds to the duration of pre-erythrocytic schizogony in the liver till the rupture of the cysts liberating up to 40,000 cryptozoites in the blood stream. A drug which can prevent this tissue stage from reaching maturity would be a *Causal Prophylactic* in that it would prevent clinical malaria though not infection. Primaquin and the 8-aminoquinolines act in this way and also proguanil and pyrimethamine.

TABLE 1. Antimalarial Drugs

Chemical Constitution	Drug	Dose Maximum therapy 24 hrs.	Dose suppress· Daily	Toxic effects or Disadvantages	Malariacidal Action: Erythrocytic schizonts All species	Gametocytes MT.	Gametocytes BT.	Pre-erythrocytic MT.	Pre-erythrocytic BT.	Relapse BT. + Q.
	Quinine	2 g.	0 5 g.	Cinchonism, Amblyopia, Blackwater f.	+ +	0	0	0	0	0
9-amino-acridine	Mepacrine (Atebrin)	(0 6 g.) 0·3 g	0·1 g.	Psychoses, Skin, aplastic anaemia	+ + +	0	0	0	0	0
4-amino-quinolines	Chloroquine, Sontochin, Camoquin	0 5 g. 0 3 g base	same once weekly	Digestive, Headaches, Asthenopia	+ + +	0	0	0	0	0
8-amino-quinolines	Pamaquin	0 03 g	Not used	Abdominal	0	+	+	Not used		△
	Pentaquine	0 06 g.		Intravascular	0	+	+			+
	Isopentaquine	0 06 g		haemolysis	0	+	+			+
	Primaquine	0 03 g.		Least toxic	0	+	+			+
Biguanide	Proguanil (Paludrine)	0·6 g.	0·1 g.	Digestive, Acquired Resistance	+ +	+	+	+	△	△
Diamino-pyrimidine	Daraprim	0 05 g.	0 025 g. weekly	—	+ +	+	+	?+	+	+
Synergic action	Quinine plus Pentaquine, etc.									+ +
	Quinine plus Chloroquine, etc.									+ +

M T. = Subtertian. + = Active. △ = Little action. B.T. = Benign tertian o = Inactive. Q. = Quartan.

ANTIMALARIALS AND THEIR ACTION

We may therefore classify the antimalarials as follows:

1.	Sporozoiticidal	or True Causal Prophylactics (Not yet known)
2.	Pre-erythrozoiticidal	Causal Prophylactics (Attack-preventing)
3.	Schizonticidal	Suppressives (Attack-abolishing)
4.	Exo-erythrozoiticidal	Curative (Relapse-preventing)
5.	Gametocytocidal	Reservoir-sterilizing, in man
6.	Sporonticidal	Transmission-blocking, in mosquito

It is easier however to consider antimalarials in groups according to their chemical constitution, and to mention the parts of the spectrum of activity which each exhibits. (See Table 1.)

The Cinchona alkaloids. Quinine is the most important and was isolated from the bark in 1820. Its formula was described in 1907 and it was synthesized in 1944. Quinine consists of a quinoline ring (6-methoxy quinoline) linked by a carbinol group to a quinucleidine nucleus. It is quickly absorbed and rapidly eliminated but 1 *gm. daily, in divided doses, is* sufficient to maintain a serum concentration of from 5 to 9 mgm./L. It is a sure and rapid schizonticide. In daily doses of 1 to 1·5 gm. for three weeks it is curative for *P. falciparum* but not for the relapsing malarias.

It is not a causal prophylactic but in daily doses of 0·3 to 0·5 gm. it acts as an efficient suppressive of moderate and occasional infections such as those to which non-immune civilians may be exposed when residing in an endemic area. If the daily dose is regular, and doubled at the slightest premonitory sign of a clinical attack, satisfactory protection is afforded. But under the rigours of military operations with repeated heavy infections, chronic malaria may develop, despite daily quinine prophylaxis, with the consequent risk of blackwater fever. There is some evidence to show, however, that a weekly or twice-weekly dose of quinine to native infants in welfare clinics in tropical Africa may be the ideal method of preventing severe fever without completely interfering with the acquisition of immunity, which depends on some degree of parasitaemia for its develop-

crine nucleus from which had been removed the benzene ring carrying the methoxyl group, or in other words a modified quinoline ring with the same basic side chain as in pamaquin. Chloroquine was soon followed by the variants named sontochin, camoquin (amodiaquin) and plaquenil.

These are the most rapidly-acting schizonticides known and a complete cure of *P. falciparum* malaria can often be effected by as little as 1·5 gm. of chloroquine in three days or of only 1 gm. of camoquin in 24 hours. As a prophylactic, good suppression is achieved either by 0·3 gm. weekly or by 0·6 gm. fortnightly. In treatment it is best to give a loading dose in order to get a 10 mgm./K. serum concentration for there is preliminary selective concentration in the liver. Nevertheless, none of the 4-aminoquinolines act directly on the pre- or exo-erythrocytic liver phases, so they are not causal prophylactics though in combination with the 8-aminoquinolines they are effective in relapse control.

Biguanides and Pyrimethamine. In the spate of activity during World War II to find new substitutes for the almost unobtainable quinine, 14,000 drugs were studied in the United States and summarized by Wiselogle in 1945. Most attention was paid to the well-known quinoline and acridine derivatives, but British workers Curd, Davey and Rose (1945) turned their attention to drugs resembling the pyrimidines because these are found in nature in the chromatin of cells, and certain derivatives, the sulphonamides, were known to be effective in malaria. Proguanil (chlorphenyl-isopropyl biguanide) was the first reward of their efforts and has recently been shown to owe its activity to a metabolite which is almost identical with pyrimethamine discovered by Hitchings (1952) and described as 2-4-diaminopyrimidine.

Both paludrine (proguanil) and daraprim (pyrimethamine) are only slowly schizonticidal, and are not therefore the drugs of choice in the clinical attack. But a daily dose of 0 1 gm. of paludrine or a *weekly* dose of as little as 25 mgm. of daraprim are excellent prophylactics. They have the additional property of rendering gametocytes, which may be ingested up to a week after the dosing of a human carrier, incapable of developing to

ment. It supplements the natural refractoriness of breast-fed infants due to the deficiency in milk of para-amino-benzoic acid. Quinine is gametocidal to *P. vivax* when used synergically with the 8-aminoquinolines.

The 8-aminoquinolines. Pamaquin was first synthesized in 1926 and consists of the quinoline ring as in quinine, with an open basic side chain identical to that which is found in mepacrine and chloroquine (*vide infra*). There followed the variants of pentaquine and iso-pentaquine, but primaquine is the latest, the least toxic and the most active. Primaquine has an action on pre-erythrocytic liver forms and can be used as a causal prophylactic though the other members are too toxic for continued administration. It also acts on the exo-erythrocytic liver stage in the relapsing malarias and is therefore curative, especially when combined with quinine or chloroquine. For this purpose 15 mgm. daily with 1 gm. of quinine or 0·3 gm. chloroquine is sufficient if continued for 14 days. Primaquine is also strongly gametocidal, rendering the gametocytes non-infectious to mosquitoes in 24 hours and banishing them from the circulation in three days.

The 9-amino acridines. Mepacrine was synthesized in 1932 and consists of a quinoline nucleus to which is fused a benzene ring to give the acridine nucleus. To this is added the same open side chain as in pamaquin (dialkylamino dialkylamino).

The action of mepacrine is exclusively schizonticidal, but it acts on a younger stage of the trophozoite than does quinine. But the necessary serum concentration of 20 mgm./L. is only slowly built up so that a loading dose is necessary, 0·6 gm. in the first 24 hours, and this is maintained by 0·1 gm. daily. Partly for this reason it is a surer suppressive than quinine and will prevent the overt attack and cure *P. falciparum* infection although *P. vivax* will become patent about 30 days after cessation of the daily prophylactic dose, in those infected with this parasite. Its value as a prophylactic under war conditions was first proved by Field in Malaya in 1939 and fully investigated and confirmed by Fairley (1945) and colleagues at the Land Headquarters Medical Research Unit at Cairns in Australia.

The 4-aminoquinolines were synthesized in 1934, using a mepa-

the gametocytes which are formed two or three weeks later are also devoid of pigment but are nevertheless viable in mosquitoes and produce sporozoites infective to man.

It appears therefore that quinine and mepacrine, and presumably the 4-aminoquinolines also, exert a deleterious action on growing forms of the parasite which are metabolizing haemoglobin. It is noteworthy that both gametocytes and sporozoites which are in a resting phase are untouched by these drugs, so that we may conclude that anabolic rather than catabolic processes are blocked by schizonticidal drugs.

Pamaquin, on the other hand, acts as a general protoplasmic poison and takes about five days to slow down schizogony when given in the safe dose of 0 03 gm. a day. A single dose however causes gametocytes to lose their infectivity to mosquitoes in 42 hours. Proguanil does not appear to interfere with the growth of schizonts, but like pyrimethamine prevents proper chromatin synthesis and blocks cellular division. A failure of mitosis may result from their anti-folinic acid action which prevents chromatin from shedding its ribonucleic acid into the cytoplasm. The theory is that this renders it 'sticky' and unable to separate.

The task of correlating chemical constitution with plasmodicidal action is more difficult. It is presumed that the function of the side chain is primarily concerned with the transport and distribution of the drug in the host and its penetration into the red cells, or the parasites, up to the particular point where the toxic action of the other part of the drug can be exerted. Plasmodicidal action is probably the function of the quinoline, acridine or pyrimidine nucleus. The process can be likened to a confidence trick in which a contact man is employed to effect the introduction of a crook who can thus get his bogus bonds accepted by an unsuspecting victim. The substitute metabolite must be a perfect fit in the metabolic chain otherwise it may be rejected or so changed that resistance develops.

Resistance

It has long been known that various natural strains of malaria parasites, *P. falciparum* notably, show very different responses to a given drug. This is a constant factor and reflects differences

mature oocysts in the mosquito's stomach. Thus although they are not strictly speaking gametocytocides they can completely block malarial transmission. Daraprim especially is opening up the possibility of having, even in a hyperendemic area, anophelism without malaria, for the weekly administration of a few milligrams to every member of the community will reduce the parasite rate to zero and bring transmission to a standstill.

How Antimalarials Act

The effectiveness of any antimalarial drug is dependent on t interaction of the parasite, the host and the drug.

The four species of malaria parasite show differing respo and different strains of *P. falciparum*, especially, vary widely virulence quite apart from a possible acquisition of di resistance. In the host, the degree of natural or acquired i munity and the condition of the reticuloendothelial system well as the general nutrition may influence chemotherapeu response. Food deficiency states may actually prolong the p patent period so that clinical malaria only develops wl famine relief has begun to operate. Similarly Hawking (19 has shown that the failure of malaria to develop in milk- animals first reported by Maegraith *et al.* (1952) is reversed the addition of para-aminobenzoate to the diet. Allison (19 has shown that the abnormal erythrocytes of individuals w the sickle-cell trait are less easily parasitized by *P. falcipa* than are normal red cells.

The dose, the manner of its administration and absorpti and the rate of its elimination, degradation or concentratio different tissues of the body all affect plasmodicidal acti Finally biological interference of two species in a double in tion may vary the picture, as also the stage of growth of parasite when confronted with the drug.

Observable morphological changes in parasites from patie undergoing quinine treatment show that there is considera delay in the growth of trophozoites which tend to lose their p ment. No true schizonts are formed. Using mepacrine sim changes are seen earlier, and at 10 hours some half-gro trophozoites have actually extruded their pigment Moreo

Whether? Is your treatment really necessary? This may be a pertinent question among native children and older residents in hyperendemic areas who have to buy their tolerance at the price of some fever. It becomes then a question of price control, of protecting them from paying too high a price.

When? As a rule treatment, especially of non-immunes, should be started as soon as a real suspicion of malaria exists. Delay is dangerous and a blood slide should be taken but treatment should not await the result. Moreover a negative thin smear may not rule out impending cerebral malaria.

What drug and what dose? can be learnt from the table which summarises the situation to date.

Malaria Control is a vast subject outside the scope of this lecture, but the accompanying diagram illustrates the bare outlines of the problem and stresses the role which the antimalarial drugs may play in malaria eradication.

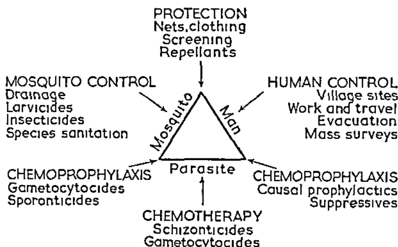


Fig. [illegible]

of virulence. But the term 'strain resistance' is given to a change in sensitivity in a given strain after exposure to a given drug. This is not found with quinine and the other schizonticides, but has been noticeable both in avian and human malaria treated with proguanil. It is manifested by a diminution of schizonticidal activity and can generally be explained by insufficient or irregular dosage. Strain resistance is conserved through the mosquito, but it does not involve a diminution of toxicity to the pre-erythrocytic stage in the liver. Strains which have become resistant to paludrine remain so for the similar daraprim also but not to the other chemically different drugs.

Application to Treatment and Control of Malaria

With such complete knowledge of the malaria parasite and with such a variety of drugs acting in different ways and with various advantages, or unpleasant side effects, we need to ask a number of preliminary questions before prescribing any treatment.

Who? Who is the patient? Is it a non-immune person suffering from the first attack of M.T. malaria, caught possibly during an air voyage through an endemic area? If so, schizonticides like chloroquine, mepacrine or quinine are indicated. Or is it a settler who has had a number of previous attacks and may have an enlarged spleen when quinine might precipitate blackwater? Or is it an indigenous inhabitant born and bred with his malaria who only needs a dose or two of any antimalarial to tide him over?

Which? Which of the four human parasites is concerned?

P. falciparum of malignant tertian fever (M.T.).

P. vivax of benign tertian fever (B.T.) with a tendency to relapses for up to four years.

P. malariae of quartan fever (Q) and its relapses up to 30 years.

P. ovale causing a mild self-limiting infection.

Why? The choice of drug and the course prescribed depends on why the patient is being treated. This may be for:

(*a*) Control of symptoms by schizonticides,

(*b*) Anti-relapse by 8-aminoquinolines,

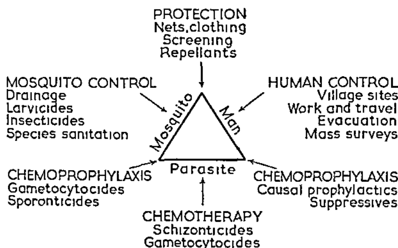
(*c*) Prophylaxis by Paludrine, Daraprim or Chloroquine.

Whether? Is your treatment really necessary? This may be a pertinent question among native children and older residents in hyperendemic areas who have to buy their tolerance at the price of some fever. It becomes then a question of price control, of protecting them from paying too high a price.

When? As a rule treatment, especially of non-immunes, should be started as soon as a real suspicion of malaria exists. Delay is dangerous and a blood slide should be taken but treatment should not await the result. Moreover a negative thin smear may not rule out impending cerebral malaria.

What drug and what dose? can be learnt from the table which summarises the situation to date.

Malaria Control is a vast subject outside the scope of this lecture, but the accompanying diagram illustrates the bare outlines of the problem and stresses the role which the antimalarial drugs may play in malaria eradication.



of virulence. But the term 'strain resistance' is given to a change in sensitivity in a given strain after exposure to a given drug. This is not found with quinine and the other schizonticides, but has been noticeable both in avian and human malaria treated with proguanil. It is manifested by a diminution of schizonticidal activity and can generally be explained by insufficient or irregular dosage. Strain resistance is conserved through the mosquito, but it does not involve a diminution of toxicity to the pre-erythrocytic stage in the liver. Strains which have become resistant to paludrine remain so for the similar daraprim also but not to the other chemically different drugs.

Application to Treatment and Control of Malaria

With such complete knowledge of the malaria parasite and with such a variety of drugs acting in different ways and with various advantages, or unpleasant side effects, we need to ask a number of preliminary questions before prescribing any treatment.

Who? Who is the patient? Is it a non-immune person suffering from the first attack of M.T. malaria, caught possibly during an air voyage through an endemic area? If so, schizonticides like chloroquine, mepacrine or quinine are indicated. Or is it a settler who has had a number of previous attacks and may have an enlarged spleen when quinine might precipitate blackwater? Or is it an indigenous inhabitant born and bred with his malaria who only needs a dose or two of any antimalarial to tide him over?

Which? Which of the four human parasites is concerned?

P. falciparum of malignant tertian fever (M.T.).

P. vivax of benign tertian fever (B.T.) with a tendency to relapses for up to four years.

P. malariae of quartan fever (Q) and its relapses up to 30 years.

P. ovale causing a mild self-limiting infection.

Why? The choice of drug and the course prescribed depends on why the patient is being treated. This may be for:

(*a*) Control of symptoms by schizonticides,

(*b*) Anti-relapse by 8-aminoquinolines,

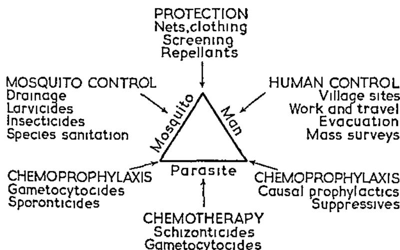
(*c*) Prophylaxis by Paludrine, Daraprim or Chloroquine.

Whether? Is your treatment really necessary? This may be a pertinent question among native children and older residents in hyperendemic areas who have to buy their tolerance at the price of some fever. It becomes then a question of price control, of protecting them from paying too high a price.

When? As a rule treatment, especially of non-immunes, should be started as soon as a real suspicion of malaria exists. Delay is dangerous and a blood slide should be taken but treatment should not await the result. Moreover a negative thin smear may not rule out impending cerebral malaria.

What drug and what dose? can be learnt from the table which summarises the situation to date.

Malaria Control is a vast subject outside the scope of this lecture, but the accompanying diagram illustrates the bare outlines of the problem and stresses the role which the antimalarial drugs may play in malaria eradication.



REFERENCES

Allison, A. C. (1954). 'Protection afforded by Sickle-cell trait against Subtertian Malarial Infection.' *Brit med. J.* **i**, 290.

Covell, G. (1953). 'Current Research towards a Global Control of Malaria.' *New. Eng. J. Med.* **249**, 125.

Curd, F. H. S, Davey, D. G. and Rose, F. L (1945). 'Studies in Synthetic Antimalarial Drugs.' *Ann. trop. Med. Parasit.* **39**, 208.

Fairley, N. H. *et al.* (1945). 'Chemotherapeutic Suppression and Prophylaxis in Malaria.' *Trans R. Soc. trop. Med. Hyg.* **38**, 311.

Fairley, N. H. (1952). 'The Chemoprophylaxis and Chemotherapy of Malaria in Man with special reference to the Life Cycle.' *Austr Ann. of Med.* **i**, 7.

Hawking, F. (1954). 'Milk, *p*-aminobenzoate and Malaria of Rats and Monkeys' *Brit med. J.* **i**, 425.

Hitchings, G H (1952). 'Daraprim as an antagonist of Folic and Folinic Acids.' *Trans. R. Soc. trop. Med Hyg.* **46**, 467.

Maegraith, B G. *et al.* (1952). 'Suppression of Malaria (*P. berghei*) by Milk.' *Brit. med. J.* **ii**, 1382.

Shortt, H E. and Garnham, P. C. C (1948). 'Exoerythrocytic Parasites in Plasmodium Cynomolgi.' *Trans. R. Soc. trop Med. Hyg* **41**, 705.

XIX

Chemotherapy of Cancer

E. BOYLAND

In some ways the title of this lecture is misleading. Firstly, much of the work which will be discussed is concerned with leucaemia and allied diseases rather than with so-called solid tumours. Secondly, there is as yet no true chemotherapy of cancer, in that the drugs used have only a temporary palliative action in relieving symptoms and prolonging life. But this does not mean that the situation will not change. I think that drugs which can effectively destroy or control cancer cells will eventually be found for the treatment of those cases of malignant disease which cannot be effectively handled by surgery or radiotherapy. The fact that a lecture with this title should appear in this series is an indication of the progress which has already been made. So much has been done in this field during the past decade that I can only deal with a few of the interesting advances.

Carcinogenic Action of Agents used in Cancer Therapy

Although this lecture is concerned with therapy of cancer, much of the work described is of value for the study of carcinogenesis and so to the possible removal of the cause of cancer. This is because many of the means used in the control of cancer at the present time are themselves carcinogenic. Thus we have an apparent paradox. Oestrogens and ionizing radiations from X-rays or radium are well-known examples of agents which have therapeutic action in cancer and are carcinogenic. Substances with both types of action include the oestrogens, urethane, nitrogen mustards and the other alkylating agents which are

considered later in this lecture. Carcinogenic hydrocarbons and amines have been used in treatment of cancer and leucaemia in man (cf. Engelbreth-Holm and Stamer, 1947). The only therapeutic agents which are not known to be carcinogenic are those which have not been tested, such as aminopterin and mercaptopurine.

Haddow (1948) has suggested that by exact and quantitative knowledge of the process of cancer induction we may possibly be able to reverse the process and so convert the cancer cells back to normal tissue. Such a process would be an ideal form of therapy. If the change of a normal cell into a malignant cell is of the nature of a mutation and the reverse change a back mutation, then to bring about the same back mutation in every cell in a tumour would be difficult by any means which we are likely to have available for some time. If any of the tumour cells remain unchanged then such cells will grow again. Other ways of treating cancer include cell destruction as is attempted by radiotherapy, and control of growth by deprivation of factors necessary for growth or neutralization of such factors with anti-metabolites.

The Specificity of Anti-Cancer Agents

One feature which is becoming clearer in the treatment of cancer is that different agents have preferential effects on different types of cancer or on cancer of specific sites (cf. Table 1). For example, oestrogens are effective in treatment of cancer of the

Table 1

Drugs	Type of disease influenced
Oestrogens	Cancer of Prostate
Oestrogens	Cancer of Breast
Androgens	Cancer of Breast
Nitrogen Mustard, TEM	Lymphadenoma
Myleran	Myeloid Leucaemia
bis-(β-Chloroethylamino-phenyl butyric acid) (CB1348)	Lymphatic Leucaemia?
ACTH and Cortisone	Acute Leucaemia
Antifolic Acid Compounds	Acute Leucaemia

prostate in man and a proportion of cases of cancer of the breast in women. Androgens have some beneficial action in a few young women with mammary cancer. Thus there is not one means for treatment of cancer but different methods or drugs are effective for specific types of cancer. This specificity is in some ways an advantage, in that specific agents are less likely to damage many normal tissues of the body. Because of this, however, potential agents for cancer therapy should be tested against a range of different tumours.

Screening Tests

Stock, Sugiura, Dobriner and Rhoads (1951) working at the Sloan Kettering Institute in New York developed a technique of assaying possible agents so that new compounds are tested on a range of different tumours or a tumour 'spectrum' as it is called. Some results (Table 2) show the type of specificity which is obtained with animal tumours. One interesting example which is shown by this data is the effect of 8-azaguanine which Kidder, Dewey, Parks and Woodside (1949) found to be an inhibitor of tumour growth when tested against the mouse adenocarcinoma EO 771. This substance inhibits only a few types of animal tumour but it has been found to have some clinical value when combined with antifolic acid compounds in the treatment of acute leucaemia.

The methods used for testing agents may depend upon measurements of the changes in growth of tumours or of duration of life of the host caused by treatment. The increase in length of life is perhaps the most practical criterion, but in the case of animals with spontaneous tumours, the expected survival time of untreated animals is so variable that any increase is [illegible]

cytology can help in indicating the possible mechanism of drug action.

Once a substance has been found to be effective in animal tumours it must be tried clinically to find out what type of human disease it is likely to benefit and secondly to determine

TABLE 2 Inhibiting Effect of Compounds in a Spectrum of Mouse and Rat Tumours
(From Stock, 1950)

Compound	Dose mg/kg/day	Mouse tumours						Rat tumours		
		Crocker Sarcoma 180	Lewis Sarcoma T241	Adenocarcinoma EO 771	Harding-Passey Melanoma	Wagner Oestrogenic Sarcoma	Patterson Lymphosarcoma	Flexner-Jobling Carcinoma	Sarcoma 39	Walker Carcinosarcoma 256
Methyl *bis* (β-chlorethyl) amine. HCl	0·5	–	–	–	–	± to +	–	++	+++	++
3-*Bis* (β-chlorethyl) aminomethyl-4-methoxy-methyl-5-hydroxy-6-methyl pyridine. 2HCl.	5·0	±	–	±	–	++	±	+++	+++	+++
8-Azaguanine	75	–	–	++	–	–	–	–	–	–
2, 6-Diaminopurine	60	–	–	–	–	–	–	–	–	–
Triethylene melamine	0·25	±	–	–	–	+	– to ±	–	±	–
4-Amino pteroyl-glutamic acid	0·25	++	–	+	±	–	++	–	+++	+++

Grading of Tumour Inhibition
– Growth more than 3/4 the average diameter of the controls
± Growth from 1/2 to 3/4 the average diameter of the controls
+ Growth from 1/4 to 1/2 the average diameter of the controls.
++ No growth or growth to 1/4 average diameter of the controls.
+++ Destruction of tumour.

the most effective method and schedule of administering the drug.

Earlier work on compounds used in treatment of cancer has been discussed by Haddow (1948), Hohl and Schinz (1949), Karnofsky, Burchenal and Escher (1950) and Gellhorn (1953). In this lecture only a few types of treatment of cancer or leucaemia are discussed. There is not space to consider the use of sodium arsenite, benzene, colchicine, viruses, haemorrhagic agents and other agents many of which are used in medicine at the present time and any of which might form a point of departure for future investigations.

Therapy by Surgical Means

A method of therapy which might be described as negative chemotherapy, in which the organ producing hormones necessary for the growth of the tumour is removed, is effective in a proportion of cases of some types of cancer. The first example of this was ovariectomy used successfully in treatment of breast cancer by Sir George Beatson in 1896. This operation was abandoned for a time but is now used with success in a proportion of cases. The operation of orchidectomy for treatment of cancer of the prostate is analogous to that of ovariectomy for breast cancer. It, too, was tried at the end of the nineteenth century but was introduced on sound experimental and clinical observations by Dr. Huggins of Chicago in 1940. In the first series beneficial results were obtained in 31 out of 45 patients treated. The procedures of ovariectomy and orchidectomy are probably effective only with cancers which are hormone-dependent.

Huggins has more recently shown that some patients with tumours of the prostate or breast do not respond to castration although the tumours are probably hormone-dependent. With the idea that tumour growth in these cases might depend on hormones produced by the adrenal glands, these glands were removed and health maintained by continued administration of cortisone. In a number of patients prolonged beneficial results have been obtained by this procedure.

Hypophysectomy is another form of treatment depending on hormone deprivation. This has been suggested for treatment of

cancer of the adrenal, prostate or breast when the tumours are dependent on the presence of pituitary hormones. This heroic treatment is being investigated and may prove to be of value.

Hormones

As improvement in patients suffering from some forms of cancer was obtained by removal of organs which would change the hormone balance, attempts to change this balance were made by administering sex hormones. The treatment of cancer of the prostate with oestrogens introduced by Herrold (1941) (see also Haddow, Watkinson and Paterson, 1944) is still the most effective form of chemotherapy of cancer. The mechanism of action of oestrogens in treatment of prostatic cancer is still uncertain. The effect may be due to neutralization of androgens, depression of the pituitary or other glands or a direct action on the tumour tissue.

Oestrogens have also produced beneficial effects in some women with cancer of the breast. Thus Haddow *et al.* (1944) observed temporary retardation in a proportion of patients suffering from breast cancer treated with synthetic oestrogens. These workers had been led to test such compounds clinically because the synthetic oestrogen triphenylchloroethylene inhibited the growth of the Walker carcinoma in rats. The clinical observations have been frequently repeated and it has been found that post-menopausal patients are more likely to benefit from this treatment than are younger women. The mechanism of action in this case is unknown but it may be an indirect action through the pituitary or ovaries.

Some patients with mammary cancer benefit from androgen treatment (cf. Galton, 1950). In this there would appear to be a rational basis for the treatment as the androgen might neutralize the action of endogenous oestrogens and so have an effect analogous to that of ovariectomy. The treatment only produces benefit in about one-fifth of the cases treated (see Table 3), and these cases cannot be pre-selected by clinical means. On the whole, however, androgen therapy is more likely to be effective in younger women who have bone metastases.

Within the last few years the adrenal hormone, cortisone and

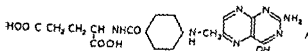
the adrenocortical hormone (ACTH) have been widely used with limited success in acute leucaemias, particularly in children. More than half of the children generally respond favourably but the response is usually of only short duration. When the disease is resistant to these hormones it may, however, still respond to medication with antimetabolites.

ANTIMETABOLITES

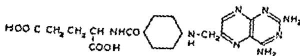
Substances which have a structural resemblance to vitamins, co-enzymes or normal intermediate metabolic products, but differ sufficiently to act as competitive inhibitors of essential metabolic processes, are known as antimetabolites. Thus ephedrine is presumably an antimetabolite and competitive inhibitor in adrenaline metabolism. Many such substances have been tried in chemotherapy of cancer. Thus sodium malonate, a competitive inhibitor of succinate oxidation, was found to reduce the growth of mammary tumours in mice (Boyland, 1940), but on clinical trial in cases of mammary cancer it was relatively ineffective.

The first antimetabolites to find successful application in the treatment of leucaemia were the anti-folic acid compounds aminopterin and amethopterin (cf. Figure 1). Aminopterin, which only differs in structure from the vitamin folic acid in that a hydroxyl group is replaced by an amino group, is a very toxic substance but both this and its less toxic methyl derivative amethopterin have produced beneficial results, including prolongation of life, in a proportion of children suffering from acute leucaemia. The leucaemias invariably become resistant to these drugs and some mouse leucaemias became aminopterin-dependent.

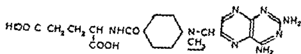
A number of purine analogues which might act as competitors have been synthesized and tested in biological systems and some of these inhibit the growth of animal tumours. One of these, azaguanine (Figure 2), an analogue and antagonist of guanine, has been discussed. Others are 2 6-diaminopurine and 6-mercaptopurine. The latter has found application in treatment of acute leucaemia in combination with anti-folic acid derivatives.



Folic Acid

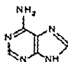


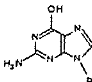
Aminopterin



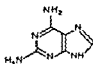
Amethopterin

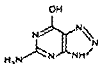
FIG. 1





Adenine *Guanine*





2 6 Diaminopurine *Azaguanine*

FIG. 2

Amino-acid derivatives such as thienylalanine and 4-*bis*-(choroethyl)-aminophenyl alanine which are related to phenylalanine and azaserine which is an antibiotic have also been found to be tumour inhibitors. Carbohydrate antagonists have also been investigated for cancer therapy. 2-Desoxyglucose inhibits the glycolysis of tumours and mitosis in tissue cultures but does not appear to influence the growth of cancer in animals.

Urethane was found by Paterson, Ap Thomas, Haddow and Watkinson (1946) to be a useful drug in the treatment of myeloid leucaemia. It has also been shown to induce tumours, mutations and chromosome breaks. The induction of chromosome damage in tumours can be partially neutralized by administration of thymine (Boyland and Koller, 1954), indicating that urethane may be an antimetabolite.

Antibiotics

Following the successes which have been achieved in the search for bacteriostatic agents in cultures of micro-organisms such cultures have been extensively examined for anti-cancer agents, particularly in the United States. Active substances of this type are called antibiotics and are often analogues or derivatives of natural metabolites and in some cases probably act as antimetabolites.

Workers in the Parke Davis Laboratories in Detroit found that culture fluids from an unidentified Streptomyces inhibited the growth of the Crocker Sarcoma 180 in mice and also the growth of the yeast *Kloeckera brevis*. The active principle has been isolated and shown to be O-diazoacetyl-L-serine

$$N_2C\overset{\overset{O}{\mathbin{/\!/}}}{C}-O-CH_2\overset{\overset{NH_2}{|}}{C}H-COOH.$$

Diazoacetyl derivatives usually have very high chemical reactivity so that it is remarkable that such a compound should occur naturally. Stock, Reilly, Brockley, Clarke and Rhoads (1954) found that the *laevo* isomer was about sixteen times [illegible]

[illegible]

[illegible] amino acid serine is modified by substitution with the chemically reactive diazoacetyl group. This diazoacetyl group could

alkylate or esterify acid groups and so this substance resembles the nitrogen-mustard derivatives (which are discussed later). It is a natural amino-acid derivative with a sting in the tail and so resembles the type of substance which Ehrlich recommended many years ago for 'Eisenbahn' therapy. In this therapy a substance with affinity for a tissue is combined with a toxic grouping which will destroy the target tissue. In azaserine, serine is possibly the train which carries the diazo residue to the tumour cells.

A series of coloured antibiotics, Actinomycins, have been isolated from Actinomycetes. One of these, Actinomycin C, inhibits the growth of animal tumours and of other tissues with dividing cells, particularly the lymphatic system, spleen and intestine. In clinical trials Actinomycin C has produced improvement in the condition of patients with Hodgkin's disease (Schulte, 1952). The molecule of Actinomycin C contains a nitrogenous pigment and a polypeptide (Brockmann, 1954). The polypeptide portion of the molecule is made up of six amino acids including the unusual amino acid sarcosine.

Radioactive Isotopes

Therapy with radioactive isotopes is another example of 'Eisenbahn' therapy where some chemical substance by virtue of its biological properties carried radioactive atoms to a specific tissue. Radioactive iodine ^{131}I has been used in this way in the treatment of cancer of the thyroid where the iodine becomes localized. This method produces improvement in a small proportion of the patients treated. Radioactive phosphorus is of value in therapy of polycythemia vera but has been somewhat disappointing as a means of treating other forms of cancer or leucaemia.

Nitrogen-mustard Derivatives

The group of substances which are fundamentally derived from nitrogen mustard have been described as radiomimetic agents on account of the similarity of their actions to those of ionizing radiations. The introduction of nitrogen mustard into therapeutics was a result of research in chemical warfare during the last war. Recognition of the similarity of the blisters produced

by mustard gas (di-2-chloroethyl sulphide) and X-rays led to the discoveries by Auerbach and Robson (1947) that mustard gas and nitrogen mustard are mutagenic and by Darlington and Koller (1947) that they produce chromosome damage similar to that caused by irradiation.

Experiments on animals and investigation of men accidentally exposed to mustard gas or nitrogen mustards showed that these substances had a leucopenic action. Following this observation, Goodman, Wintrobe, Dameshek, Goodman, Gilman and McLennan (1946) tried the nitrogen mustards HN2 or *bis*-(2-chloroethyl)-methylamine and HN3 or *tris*-(2-chloroethyl)-amine in a number of patients with leucaemia and related diseases. Improvement in condition was obtained in most of the patients suffering from Hodgkin's disease treated with these drugs. The nitrogen mustard HN2 which now has the official names chlormethine and mustine has been used with some success in treatment of lymphadenopathies in clinics all over the world. Although the first treatment usually produces a marked beneficial effect, the disease generally recurs and patients develop resistance to treatment. A few cases of Hodgkin's disease which have become radio-resistant following radiotherapy still respond to nitrogen-mustard therapy. The disadvantages of the aliphatic nitrogen mustards are that they must be given intravenously, preferably in large volumes of saline, that they cause nausea and vomiting possibly analogous to radiation sickness and that the disease eventually becomes resistant to the drug.

Since 1946 many investigators have studied compounds related to the nitrogen mustards but it is doubtful if any substances produce better effects than HN2 and HN3 in the treatment of Hodgkin's disease. In the Sloan Kettering Institute and in London a number of aliphatic chloroethylamines were synthesized and tested. One of these substances, *tetra*-(2-chloroethyl)-ethylenediamine $\begin{array}{l} CH_2N(CH_2CH_2Cl)_2 \\ | \\ CH_2N(CH_2CH_2Cl)_2 \end{array}$ was effective in inhibition of growth of animal tumours but on clinical trial had no advantages over the original compounds. These investiga-

tions showed clearly that active compounds usually contained more than one chemically reactive chloroethyl group (cf. Boyland, Clegg, Koller, Rhoden and Warwick, 1948).

Aromatic Nitrogen Mustards

In extension of research on carcinogenic and growth inhibitory derivatives of 4-aminostilbene, a compound, 4-N:N-*bis*-(2-chloroethyl)-aminostilbene (CB 1087), was prepared, examined and found to have growth inhibitory properties when tested against the Walker carcinoma. Following this observation a number of aromatic *bis*-(2-chloroethyl)-amines were synthesized and examined for biological activity (Haddow, Kon and Ross, 1948). The N:N-*bis*-(2-chloroethyl) derivatives of aniline, *o*-, *m*-, and *p*-anisidine, β-naphthylamine, and *p*-aminobenzoic acid all have marked inhibitory action, and have prolonged the lives of tumour-bearing animals. These substances have some of the properties of nitrogen mustards, but they are much less toxic and are effective when administered by mouth.

In this series of aromatic nitrogen mustards, Haddow and Ross (see Ross, 1953) have shown that the biological activity as measured by the ability to inhibit the growth of the Walker carcinoma in rats is paralleled, to some extent, by the chemical reactivity of the compounds. Compounds of this type which react with water or with acid groups at less than a certain rate are biologically inactive. Ogston (1948) had shown that mustard gas and the nitrogen-mustard compounds react readily with anions or acid groups such as the carboxyl groups of proteins and the phosphoric-acid residues of nucleic acids in neutral solutions to give complex esters. These substances are therefore alkylating agents.

The first of the aromatic nitrogen-mustard compounds to have wide clinical use (cf. Galton, 1951) was 2-*bis*-(2-chloroethyl)-naphthylamine (CB 1048) which like chlormethine produced some improvement in patients suffering from lymphosarcoma and lymphadenoma. It had great advantages in that it was less toxic, could be given by mouth and did not produce nausea. The drug, however, is much less active than HN2 but is still used for some cases of Hodgkin's disease.

Since the introduction of 2-*bis*-(2-chloroethyl)-naphthylamine into clinical use many more aromatic nitrogen mustards have been synthesized and tested on animal tumours by Professor Haddow. Of the series of the *bis*-(2-chloroethyl) phenyl fatty acid derivatives the compound with $n = 3$ or *bis*(2-chloroethyl) phenylbutyric acid (CB 1348) (Figure 3) is a powerful cytotoxic agent and inhibits the growth of animal tumours. It has been investigated in animals by Professor Haddow and Dr. Elson and in patients by Dr. Galton. In animals it shows a preferential action on the lymphocytic elements rather than on the myeloid cells. This phenylbutyric-acid derivative seems to be a better drug for Hodgkin's disease than is CB 1048 and it has had a beneficial effect in a few cases of lymphatic leucaemia in which it has been tried.

Bergel and Stock (1954) have synthesized a new type of nitrogen mustard in which the *bis*-(chloroethyl)amino group is introduced into the molecule of the amino acid phenylalanine to make corresponding *p*-*bis*-(2-chloroethyl)aminophenyl alanine. This compound CB 3007 (Figure 3), a carboxy aromatic nitrogen mustard, like CB 1348, acts preferentially on lymphocytic cells in animals.

Other Substances Derived from Nitrogen Mustard

All the substances of the nitrogen mustard type cause chromosome damage including chromosome breakage similar to (but not always identical with) that caused by X-rays and other radiations. Investigations in this and other series showed clearly that not only was a minimal rate of chemical reactivity necessary for biological activity but substances with two or more groups generally were the most effective. For this reason Goldacre, Loveless and Ross (1949) advanced the hypothesis that the active compounds were able to cross-link some constituent of chromosomes and so inhibit cell mitosis and cause chromosome damage. This hypothesis of the action of these drugs has still not been proved, but it has been a valuable idea which has stimulated work and development of the subject.

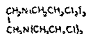
In the first place it was suggested that the known cross-linking agents used in the textile industry might be of value as chemo-

therapeutic agents for cancer. Diepoxides are a class of chemical compounds used in the textile industry for this purpose and many of these including the simplest, butadiene diepoxide, were found to be potent inhibitors of the growth tumours in animals. Butadiene diepoxide was tried clinically but was found to cause renal damage in relatively small doses.

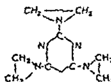
Because of the suggested hypothesis and because aliphatic nitrogen mustards are known to react through the intermediate formation of ethylenimines another known cross-linking agent 2:4:6-triethyleneimino-s-triazine or triethylene melamine was investigated simultaneously in London (Haddow, 1950), Manchester (Rose, Hendry and Walpole, 1950), and New York (Burchenal, Crossley, Stock and Rhoads, 1950), and in all these centres it was found to inhibit tumour growth. This drug also known as TEM is now widely used in America for chemotherapy of lymphadenopathies. It has the advantages that it can be administered by mouth, and that by giving divided doses, nausea and vomiting can be avoided. It is, however, less certain in its action than chlormethine and in some cases has produced very distressing reactions of bone-marrow depression.

Myleran

Another line of investigation in the aliphatic alkylating agents has been pursued by Haddow and Timmis (1953). Starting with the assumption that nitrogen mustards are alkylating agents, compounds were made in which the active chlorine of the mustards was replaced by a *p*-toluene-sulphonic ester or tosyl group. Such compounds could react with acid radicals or with amino groups. A simple tosyl analogue (435, Figure 3) of *bis-(chloroethyl) aniline (CB 1074) was found to inhibit the* growth of the Walker carcinoma in rats. Compounds of this type had less destructive action on the bone marrow and the corresponding methanesulphonoxy or mesyl derivatives, had greater chemical reactivity and also greater biological activity than the corresponding tosyl compounds. Although *bis* chloroethyl ether $ClCH_2CH_2OCH_2CH_2Cl$ was inactive, the *bis*-mesyl diethylene glycol $CH_3SO_2OCH_2CH_2O_2CH_2CH_2O\text{-}SO_2CH_3$ (CB 1020) was an inhibitor.



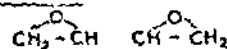
Tetrachloroethylethylenediamine SK 107
No advantage over HN 2



Triethylenemelamine (TEM)
Used for Hodgkin's Disease

Elaboration and duplication of molecule (Burchenal, Karnofsky and Rhoads)

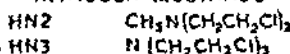
Demonstration of reaction through ethyleneimine form (Bayland Rydon)



Butadiene diepoxide too toxic for clinical use

Cross-linking Hypothesis (Goldacre, Loveless & Ross)

NITROGEN MUSTARDS



Used for Hodgkin's Disease

Combination with aminostilbene (Haddow, Kon and Ross)

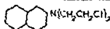




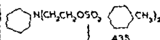
-N(CHCHCl)$_2$ bis chloroethylaniline - simplest aromatic nitrogen mustard

Variation of molecule (Haddow and Ross)

Replacement of Cl by Tosylether (Haddow and Timmis)



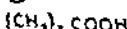
CB1048 used in cases of Hodgkin's Disease



435

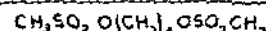




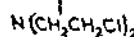


CB1348 specific for lymphatic leucaemia

Simplification of molecule
$CH_3SO_2O(CH_2)_nOSO_2CH_3$

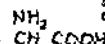


Myleran
specific for myeloid leukaemia
(Galton)



(Bergel and Stock)
bis (chloroethyl) aminophenylalanine





Substances enclosed in rectangles have been used clinically

Fig. 3

therapeutic agents for cancer. Diepoxides are a class of chemical compounds used in the textile industry for this purpose and many of these including the simplest, butadiene diepoxide, were found to be potent inhibitors of the growth tumours in animals. Butadiene diepoxide was tried clinically but was found to cause renal damage in relatively small doses.

Because of the suggested hypothesis and because aliphatic nitrogen mustards are known to react through the intermediate formation of ethylenimines another known cross-linking agent 2:4:6-triethyleneimino-s-triazine or triethylene melamine was investigated simultaneously in London (Haddow, 1950), Manchester (Rose, Hendry and Walpole, 1950), and New York (Burchenal, Crossley, Stock and Rhoads, 1950), and in all these centres it was found to inhibit tumour growth. This drug also known as TEM is now widely used in America for chemotherapy of lymphadenopathies. It has the advantages that it can be administered by mouth, and that by giving divided doses, nausea and vomiting can be avoided. It is, however, less certain in its action than chlormethine and in some cases has produced very distressing reactions of bone-marrow depression.

MYLERAN

Another line of investigation in the aliphatic alkylating agents has been pursued by Haddow and Timmis (1953). Starting with the assumption that nitrogen mustards are alkylating agents, compounds were made in which the active chlorine of the mustards was replaced by a *p*-toluene-sulphonic ester or tosyl group. Such compounds could react with acid radicals or with amino groups. A simple tosyl analogue (435, Figure 3) of *bis*-(chloroethyl) aniline (CB 1074) was found to inhibit the growth of the Walker carcinoma in rats. Compounds of this type had less destructive action on the bone marrow and the corresponding methanesulphonoxy or mesyl derivatives, had greater chemical reactivity and also greater biological activity than the corresponding tosyl compounds. Although *bis* chloroethyl ether $ClCH_2CH_2OCH_2CH_2Cl$ was inactive, the *bis*-mesyl diethylene glycol $CH_3SO_2OCH_2CH_2O_2CH_2CH_2O\text{-}SO_2CH_3$ (CB 1020) was an inhibitor.

neutropenic action and produced delayed aplasia of bone marrow. When it was tried in repeated small daily doses on a group of patients suffering from myeloid leucaemia, considerable relief and improvement of condition was observed (Galton, 1953). The changes usually included an increase in the haemoglobin concentration, a decrease in the total white-cell count, and particularly in the immature forms, with little effect on the platelets. Myleran now seems to be established as the best drug for the treatment of chronic myeloid leucaemia: it is more certain in its action and has less side effects than arsenite or urethane.

Elson (1953) has carried out investigations on the comparative effect of X-rays, aromatic nitrogen mustards and Myleran on the blood picture and the growth of rats. When rats are given a sub-lethal dose of *bis*-(chloroethyl)-aminophenylbutyric acid (CB 1348) a rapid fall in the number of circulating lymphocytes occurs followed by a temporary rise in neutrophils and undifferentiated cells some days later (see Figure 4). In contrast to this a sub-lethal dose of Myleran (15 mg./kg.) causes a gradual fall in the circulating neutrophils, with only slight effect on lymphocytes. There is also a fall in the platelet count, and multiple haemorrhages (particularly in the intestinal tract) develop a week later. This dose of Myleran causes the same fall in neutrophil count as would whole-body irradiation with a dose of 400 r. Such irradiation, however, also causes an 80 per cent fall in lymphocytes. Thus, while X-rays destroy both neutrophils and lymphocytes, Myleran has a more specific action on the neutrophils and CB 1348 attacks the lymphocytes. By treatment with Myleran and CB 1348 together effects similar to those produced by X-rays are seen. This shows that by treatment with chemicals one can obtain more specific effects than can be obtained with radiation.

Drug Resistance

One of the difficulties in chemotherapy of infective diseases is the development of drug resistance, and this difficulty is even greater in cancer treatment. Most of the agents which are used for the treatment of cancer are mutagens and cancer cells tend to have less genetic stability than normal cells. Thus, in treating

Following these observations the series of dimesyl glycols $CH_3SO_2O(CH_2)_nOSO_2CH_3$ was prepared and tested in animal experiments. The compound with a chain of four carbon atoms, 1:4-dimethanesulphonoxybutane (CB 2041) was found to have maximal chemical reactivity and maximal inhibitory action on the Walker carcinoma. This butane glycol derivative, now called Myleran, was tested clinically. Short courses of relatively large doses were found to have a strong

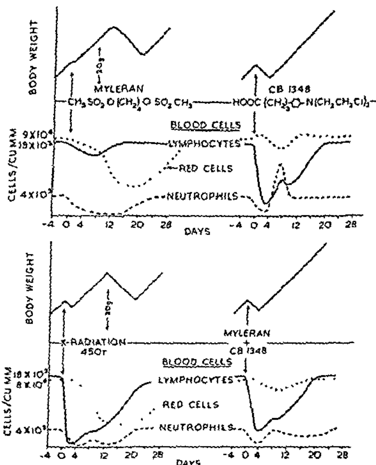


FIG. 4. Changes in body weight and in blood cells of rats treated with Myleran, with *p-bis* (chloroethyl) phenyl butyric acid, with X-rays and with the two drugs together *(from experiments of Dr. L. A. Elson)*.

In this lecture a few of the aspects of research on the chemotherapy of malignant diseases in which progress is being made at the present time have been discussed. That there is so much to report may be, as Volker suggested a century ago, because 'the degree to which a disease is open to attack is inversely related to

TABLE 3. Response of Patients with Breast Cancer to Different Androgen Preparations (from Galton 1950)

Preparation	Success	Moderate success	Failure	Total
Testosterone (implant)	5 (3)	7 (5)	14 (6)	26
Methyltestosterone (sublingual)	12 (5)	8 (2)	20 (7)	40
Testosterone propionate (intramuscular)	2 (1)	4 (3)	3 (1)	9
Combined—Testosterone and Methyltestosterone	2 (1)	1 (1)	—	3
Testosterone and Testosterone propionate	0	0	0	0
Methyltestosterone and Testosterone propionate	1	3 (2)	2	6
Testosterone, Methyltestosterone and Testosterone propionate	1 (1)	0	0	1

Figures in parentheses indicate cases with skeletal involvement
Cases treated concurrently with X-rays omitted

a number of remedies which we possess'. Woglom said only seven years ago that 'to find a remedy for cancer is almost—not quite, but almost—as hard as finding some agent that will dissolve away the left ear say, yet leave the right ear unharmed, so slight is the difference between the cancer cell and its normal ancestor'. The progress which I have described may suggest that Woglom was perhaps unduly pessimistic.

REFERENCES

AUERBACH, C and ROBSON, J. M. (1947). *Proc roy Soc. Edin.* **62**, 271
BERGEL, F. and STOCK, J. A. (1954) *J. Chem. Soc.* 2409
BOYLAND, E (1940). *Biochem. J.* **34**, 1196
BOYLAND, E, CLEGG, J. W, KOLLER, P. C., RHODEN, E. and WARWICK, O. H (1948). *Brit. J. Cancer*, **2**, 17.

cancer one is often treating unstable cells with substances which cause mutations, so that it is not surprising that resistant cells arise. Tumours and leucaemias become resistant to irradiation or to drugs. In treatment of tuberculosis the chance of resistant organisms developing is reduced by using combined therapy such as streptomycin with *p*-aminosalicylic acid, or streptomycin with *iso*nicotinic hydrazide. By such treatment organisms resistant to streptomycin are killed by the second drug before they are present in large numbers. Such combined therapy would seem to offer special advantages in cancer therapy but the drugs used must be chosen so that they work through different mechanisms. It has been suggested that by blocking of sequential processes in cell metabolism, enhanced inhibitory effect could be obtained. Thus, by using a folic acid antagonist (aminopterin) and a purine antagonist (6-mercaptopurine) a more prolonged inhibition of tumour growth is obtained than by using each drug separately. This combination is being used in treatment of acute leucaemia.

Variation of Response

An outstanding phenomenon in chemotherapy of cancer is the great variation in the response which is observed. None of the remedies which have been discussed is effective in all cases of the type of disease in which it has beneficial action. For example although the treatment of prostatic cancer with oestrogens is perhaps the most efficacious form of chemotherapy of cancer, it is only effective in about two-thirds of the cases treated. In the treatment of mammary cancer with androgens only about one-quarter of the treated cases show improvement (see Table 3, from Galton, 1950). In both these types of cancer it is not yet possible to decide by previous examination whether a patient is likely to respond to the treatment or not.

This uncertainty as to the outcome of treatment also applies to a less extent to the chronic leucaemic condition. The most recent advances have been made in treatment of leucaemia but the knowledge and experience gained in this field will be of value in investigating the treatment of other malignant conditions.

XX

The Scientific Approach to Dermatology

R. M. B. MACKENNA

There is little doubt that many members of our profession would still be prepared courteously to express a little scepticism concerning the title of this lecture Dermatology—they would say—is largely a matter of clinical study; the practitioners of this art are addicted to giving incomprehensible names of considerable length and dubious latinity to a variety of eruptions; to the normal, non-dermatological eye many of these eruptions look remarkably similar; and, if our critics are feeling a little rancorous, they may add that it is typical of a dermatologist that, having laboriously reached an erudite diagnosis, he usually hastens to the matter of treatment, and—apparently with no feeling of anti-climax—gravely but invariably prescribes one of three simple remedies, calamine lotion, zinc cream or Lassar's paste. One can—our critics imply—save a lot of time by prescribing treatment first and adding the diagnosis to the case sheet afterwards.

There is, perhaps, a modicum of truth in these reproaches, but I hope that this lecture will show that there is a good deal more in dermatology than our critics suggest. Expressed more tersely, our difficulty has been that dermatology has taken longer than most other branches of what used to be called the 'Natural Sciences' to pass through the era of classification and nomenclature, and this greatly affects our approach to our subject. There are almost exact parallels in zoology and botany: for years the exponents of these sciences argued and wrote about species and genera and names; but nowadays biologists are much more interested in physiology, chemistry and even

BOYLAND, E. and KOLLER, P. C. (1954). *Brit. J. Cancer*, in press
BROCKMANN, H (1954). *Angewandt Chemie*, **66**, 1.
BURCHENAL, J. H., MURPHY, M. L., ELLISON, R. R., SYKES, M. P., TAN, T. C, LEONE, L. A., KARNOFSKY, D. A., CRAVER, L. F, DARGEON, H. W. and RHOADS, C. P. (1953). *Blood*, **8**, 965.
BURCHENAL, J. H., CROSSLEY, M. L., STOCK, C. C. and RHOADS, C. P., (1950) *Arch. Biochem* **26**, 321.
DARLINGTON, C. D. and KOLLER, P. C. (1947). *Heredity*, **1**, 187.
ELSON, L. A. (1953). *Ann. Rep British Empire Cancer Campaign*, **30**, 351.
ENGELBRETH-HOLM, J. and STAMER, S. (1947). Approaches to Tumor Chemotherapy. *Am Ass. Adv Sci.*, Washington D.C. 418.
GALTON, D. A. G. (1950). *Brit. J. Cancer*, **4**, 20.
GALTON, D. A. G. (1951). *Brit. J. Radiol.* **24**, 511.
GALTON, D. A G. (1953). *Lancet*, **i**, 208.
GELLHORN, A. (1953). *Cancer Res.* **13**, 205.
GOLDACRE, R. J., LOVELESS, A and ROSS, W. C. J. (1949). *Nature*, **163**, 667.
GOODMAN, L. S., WINTROBE, M M., DAMESHEK, W., GOODMAN, M. J., GILMAN, A. and MCLENNAN, M. T. (1946). *J. Amer med. Assoc.* **132**, 126.
HADDOW, A. (1948). *Brit. med. Bul* **4**, 417.
HADDOW, A. (1950). *Ann. Rep. British Empire Cancer Campaign*, **28**, 56.
HADDOW, A., KON, G. A. R and ROSS, W. C J (1948). *Nature*, **162**, 824.
HADDOW, A. and TIMMIS, G. M. (1953). *Lancet*, **i**, 207.
HADDOW, A., WATKINSON, J. M. and PATERSON, E (1944). *Brit. med. J.* **2**, 393
HERROLD, R. D. (1941). *J. Urol.* **46**, 1016.
HOHL, K and SCHINZ, H R. (1949). *Schweiz. med. Wschr.* **79**, 421.
JACQUES, J. A , STOCK, C. C. and BARCLAY, R. K (1953). *Cancer*, **6**, 828.
KARNOFSKY, D. A , BURCHENAL, J. H. and ESCHER, G C. (1950). *Med. Clin. N Amer.* **34**, 439.
KIDDER, G W , DEWEY, V. C , PARKS, R. E. and WOODSIDE, G. C. (1949). *Science*, **109**, 511.
OGSTON, A G. (1948) *Biochem. Soc Symposia*, **2**, 2.
PATERSON, E., AP THOMAS, I , HADDOW, A and WATKINSON, J. M (1946). *Lancet*, **i**, 677.
ROSE, F. L., HENDRY, J. A. and WALPOLE, A. (1950). *Nature*, **165**, 993.
ROSS, W. J C (1953) *Advances in Cancer Research*, **i**, 397
SCHULTE, G. G (1952) *Z Krebsforsch* **58**, 500.
STOCK, C. C. (1950). *Amer J Med* **8**, 658.
STOCK, C. C , REILLY, H. C , BROCKLEY, S. M., CLARKE, D. A and RHOADS, C P (1954). *Nature*, **173**, 71
STOCK, C. C., SUGIURA, K , DOBRINER, K and RHOADS, C. P. (1951) *Acta Un int. Cancr* **7**, 530.

active professional life was the initiator of the scientific approach to dermatology. He published over 600 papers (many very brilliant: a few not so commendable); and was a pioneer in histo-pathology, histo-chemistry and bacteriology. His career was illumined by a long series of investigations based on his original application of the knowledge of the physical and biological sciences of his day to the problems of dermatology. His influence on all dermatological schools has been tremendous, greater than that of Hebra whose major interest was clinical work and clinical demonstration. It was Unna who taught us to appropriate all the aids of the modern sciences of our time for the investigation and comprehension of our problems.

Unna, however, was a dermatologist whose work was chiefly addressed to dermatologists, and for many years this state of affairs largely persisted—no one except the skin specialist was very interested in the skin. It is only comparatively recently, in this country at any rate, that—perhaps stimulated by the researches of Sir Henry Dale and Sir Thomas Lewis—men who are 'scientists' and not 'clinicians' have become more and more interested in the integument, so that now in many laboratories throughout the country research—and very erudite research too—is being undertaken concerning the structure and activities of the skin in health and in disease.

This neglect of the skin—the largest organ of the body—has been due to the great difficulties which beset those who try to fathom its secrets. If the skin was thicker, investigation would be easier, but the investigator finds himself confronted by what is almost a membrane which is wide in length and breadth but —if Muchow's (1925) estimates are accepted—has an average thickness of only 2·2 mm.; the epidermis, which is a laboratory of amazing complexity synthesizing keratin, pigment, vitamin D precursor, lipids, enzymes and anti-enzymatic factors, has an average thickness of 0 12 mm.: the corium which, besides its physical properties of elasticity, resilience, and ensheathment, has functions in metabolism—collagen formation, for example —immunology and heat regulation, undertakes these and other very diverse duties all in an area that on the average is reported

psychology than they are in the disputes which made and marred reputations two generations ago. Today I find myself a member of a generation which is much more interested in how an eruption has materialized than in the classical accuracy of the terminology in which the diagnosis is written. The diagnosis, of course, has to be accurate but we are not particularly concerned whether the terms we use would be approved by Ferdinand Hebra or one of his immediate successors. For example, as the nodules of erythema nodosum fade the affected skin has a bruised appearance; classically the eruption in this stage is Dermatitis Contusiformis, but most of my generation would prefer to forget this and give the name Erythema Nodosum in which the aetiology is connoted by custom.

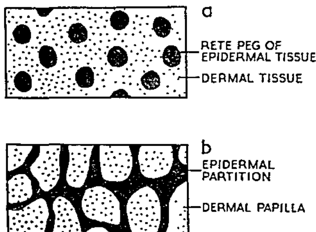
Now the reason why dermatology has lain so long in the doldrums of classification and nomenclature is because the associated sciences, particularly anatomy, physiology and biochemistry, are only now reaching phases in their development in which they are able adequately to investigate the structure and the metabolic processes of the skin; until we know more about these it is impossible fully to comprehend what may occur in disease and thus to understand not only how a rash appears but how it is that, for example, in the group known as the Lichens the primary and—classically—the persisting lesions are papules, whilst in the Herpetiform group the primary lesions are vesicles.

It follows, in my view, that, briefly, the scientific approach to dermatology is the path whereby we will discover how the skin functions in health, how it is impaired in disease and how tumours—benign and malignant—are formed. A later generation may be permitted to ask not 'how' these things happen but 'why' they occur; why, for example the skin is chosen in a given case to produce a definite symptom. But, as I. Macalpine (1953) has stated, 'why' is not at present a legitimate question to ask of the natural sciences, for as yet both the question and its answer lie in the domain of philosophy.

Basic Data

It is usually accepted that Paul Gerson Unna who was born in Hamburg in 1850 and who died in 1929 after nearly 60 years of

these have perhaps received less attention. E. H. Leach (1952) has described a technique whereby sections of skin 500μ in thickness can be cut, stained with anthracene blue and examined with a (stereoscopic) binocular microscope. By cutting sections in planes parallel to the surface he has shown that two or more separate sebaceous glands may open into the same follicle, and that there is no evidence for the concept of 'rete



[illegible] VII, [illegible] ere [illegible] in horizontal sections through the dermal papillae (from Leach, 1952).

pegs'. No longer may we regard the epidermis as having finger-like projections extending downwards to the corium, for it seems that the cone-shaped dermal papillae indent the inner surface of the epidermis. The papillae are so closely placed that only thin partitions of epithelium separate them. 'It is as though the dermal papillae were surrounded by flexible plastic cups set upside down and squeezed together until their sides touched.' In the standard, thin, vertical sections the walls of these 'cups' [illegible]

note that Malpighi himself gave a correct description, which

to be but 2 mm. in depth. The hypoderm, which is *terra incognita* at present, is said to be 0·08 mm. thick. When we consider this we can understand why physiologists and others turned from the skin to other organs more amenable to the techniques which they had elaborated or could devise. Yet, despite its thinness, knowing as we do the infinite capacity of Nature, working by a process of selective evolution to obtain a maximum of efficiency from all the essential organs of the bodies of higher vertebrates, it would have been reasonable to assume, on purely philosophical grounds, that the skin had a complexity of functions beyond those classified in the elementary text-books under headings such as integumentary, tactile, heat regulation, etc.

The skin, of course, is the largest organ of the body. Morris Leider (1949) has estimated that if by 'skin' is meant epidermis, corium, and that portion of the hypoderm which ensheathes the deepest appendages, its weight is 6 per cent of the body weight.

Here I cannot forbear referring to an illustration of a statue of the patron Saint of St. Bartholomew's Hospital (Plate VI, Figure 1);[1] this statue is in Milan Cathedral and was executed by Marco d'Agrata of Ferrara who was born about 1500. St. Bartholomew was one of the twelve apostles; according to the traditional account he suffered martyrdom by being flayed alive and then crucified head downwards at Albanopolis in Armenia or Urbanopolis in Cecilia. Here he is shown carrying his skin over his shoulders; and if his weight was 12 stone, the weight of that skin (disregarding the subcutaneous tissue which seems also to have been removed) was a fraction over 10 lb.—more than twice the weight of the liver and nearly three times that of the brain.

Anatomy and Physiology

It is interesting to note that in the University of Oxford in recent years much valuable pioneer investigation has been done in several different departments concerning the skin. The work done by Sir Rudolph Peters and his colleagues concerning Vitamin B pyrophosphate ('Co-carboxylase') and the elaboration of dimercaptopropanol is widely known, but at the moment, I wish to turn to matters concerning anatomy, for

[1] The plates referred to in this lecture will be found at the end of the book

hair follicles, myelinated parent fibres enter the follicles, and end in a profusion of filaments disposed in two planes. Encapsulated nerve endings are seen only in glabrous skin and mucous membrane. The capsule is formed of epithelial cells and connective tissue fibres and within it the parent axon pursues a tortuous course and from it springs a profusion of short, fine, naked filaments which branch repeatedly and terminate among the cells of the capsule.

In summary, therefore, all the nerve fibres which end in the skin ultimately give rise to a series of fine, naked, axoplasmic filaments.

Weddell (1953) believes that the conception of one nerve ending for each nerve fibre, forming one pathway for one perception—a concept based on the validity of the law of specific irritability—can no longer be regarded as a sound working hypothesis.

Sir Thomas Lewis (1942) coined the name 'nocifensor' for a system of nerves in the skin and mucosae which are concerned with local defence against injury. Lewis' observations suggested that there was a dorsal root efferent system of nerve fibres which formed a network just below the epidermis and which was concerned, among other things, with the spread of flare and with hyperalgesia. He argued that this system was independent of nerves subserving 'pain', for nocifensor reactions are not necessarily accompanied by that sensation. Woollard, Weddell and Harpman (1940), on the other hand, demonstrated that in an area of skin subserved by terminals from a single nerve fibre, the skin was hyperalgesic although only pain could be evoked.

Weddell believes that the nerve terminals ending in relation to the capillaries are so disposed that local responses arising from stimuli involving short segments of superficially situated neighbouring terminals can easily spread far enough in the arborization of filaments to produce a flare without necessarily extending all the way to the main fibres which they subserve. Thus these nerve terminals and the ramification of nerve filaments could be regarded as an independent physiological system. This accords with Lewis' conception of a nocifensor

has been followed in a few texts, but that the error of describing rete pegs had crept into many books by the end of the nineteenth century.

The problems of cutaneous innervation used to be dogmatically taught and easily learned. It was said that there were no nerves in the epidermis, whilst in the corium there were Pacinian corpuscles, Krause's end bulbs and various structures which subserved not only the sensations of pain, heat, pressure and so forth, but also—nicely drawn—provided decorative illustrations for the text-books. But, as Ovid wrote in the *Metamorphoses*, 'quod fuit ante, relictum est' ('that which was previously is left behind'), and G. Wedell and W. Pallie (1954) of Oxford have shown that when improved histological methods are used, employing the enzyme hyaluronidase, the nerves can be seen to terminate in the skin in a manner which so far has only been hinted at and which has not previously been demonstrated unequivocally. 'Briefly, numerous nerve fibres reach the skin after branching and ramifying in the cutaneous nerve plerus. These fibres may be myelinated or non-myelinated. In some cases, the myelinated ones lose their myelin sheaths and travel for several millimetres as non-myelinated fibres before giving rise to a profuse arborization of fine, freely ending, naked filaments.' These filaments are usually enlarged at their points of termination. Nerve terminals in the skin may be divided into three groups: first, unencapsulated; secondly, encapsulated; thirdly, endings related to hair follicles.

Numerous naked axoplasmic filaments ramify in the epidermis among the cells of *stratum germinativum*. A wide area of epidermis, which in the forearm may be a square centimetre in extent, may be supplied by an arborization of those filaments which spring from a single stem axon in the corium. The filaments do not enter the cytoplasm of the cells.

Unencapsulated nerve endings are found in rich profusion

Just below the sites where the sebaceous glands open into the

that prolonged itching is—for many—a most demoralizing symptom. In the initial stage scratching is commonly pleasurable, but the pleasure is soon lost; the patient often despises himself for having to give way to what he regards as a bestial habit. The itching subsides, or is mitigated when pain supervenes. No one, I believe, has yet estimated whether the intensity of itching must be countered by an equal intensity of pain before the attack is controlled. My impression is that if you compare the excoriations caused by scratching to relieve an itching eruption on the forearms, with the excoriations sometimes inflicted for a similar purpose on the hand, the lesions on the latter part are more closely set than on the former—as though the patient can obtain relief by much less trauma on the forearm than on the hand; does this fit in any way with Weddell's observation that naked axoplastic filaments arising from a single stem axon ramify in a much smaller area (a few square millimetres in extent) on those parts of the fingers which he has examined than on the forearm where their area of ramification [illegible]

present; and in doing this not only will we be able to relieve more human suffering, but we will undoubtedly hasten the rate of cure in many eruptions which are kept active—often for months—by friction and scratching

Metabolism

I would like now to draw attention to matters concerning the metabolic activity of the skin in health and disease. Until recently it was taught that when ultra-violet rays fall on the epidermis certain sterols in the skin are irradiated and vitamin D is formed; the individual is thus safeguarded from rickets in early childhood. Both in childhood and in later life it is probable that his resistance to tuberculosis is enhanced, for, as G. B. Dowling (1953) has shown, the vitamin whilst having no effect on the growth, viability and virulence of the tubercle bacillus nevertheless stimulates the host's specific cellular reaction to the organism. V. R. Wheatley informs me that most biochemists

system behaving independently of the sensory system of pain nerves. It is of course to be assumed that the local response is connected with changes whereby chemical substances are released which influence the state of the capillaries.

From the clinical aspect dermatologists welcome any information which increases their knowledge of cutaneous innervation; we are interested in inflammatory responses and 'flare', but patiently we wait for someone to shed new light on the problem of itching. The last important paper on this subject was probably that published by S. Rothman (1941); he showed that tickling, itching and diffuse burning pain all represent the same sensory quality and are felt in turn as the intensity of the stimulus is increased. Possibly a full consideration of Weddell's observations may 'tie up' with some of Rothman's data.

Itching may be central or peripheral. J. H. Sequeira, J. T. Ingram and R. T. Brain (1947) have suggested that the immediate cause of itching is a change in the surface tension between the prickle-layers of the epidermis.

Severe itching encountered in eczema, in dermatitis herpetiformis and lichen planus is, so many patients say, worse to endure than pain. The same remark has been made by some patients suffering from Hodgkin's disease, and by very many who have *pruritus ani* or *pruritus vulvae*.

In the present state of our knowledge I would not like to say that the aetiology of the symptom is the same in all these maladies; in the first group—eczema, dermatitis herpetiformis and lichen planus—the itch may be due solely to peripheral causes: is it solely peripheral in the early stages of psychogenic *pruritus ani* or in Hodgkin's disease?

Wittkower and Brian Russell (1953) regard itching as being an epidermal and upper dermal function, but they are careful to emphasize that dermal reactions such as angio-neurotic oedema do not itch It is difficult to understand why the lesions of urticaria itch whilst those of erythema multiforme cause more a burning than an itching sensation.

Perhaps it may be thought that I have spent too much time on this matter of itching· but, as a clinician, I would like to emphasize that it is much easier to relieve pain than itching, and

'impetigo' was giving concern to the military authorities and the average duration of stay in hospital in these cases seemed to be far too long. It was thought that when penicillin cream became available, men with 'impetigo' would be treated in their units: they would be rapidly cured and sick wastage would be greatly reduced. It was soon found that this did not happen. The sick wastage from this maladay remained about as great as ever. Twiston Davies and his colleagues (1945) investigated the clinical aspect of the matter. He found that a large proportion of these men were suffering from seborrhoeic dermatitis (see Plate VII, Figure 4). The lesions often had become infected and mimicked impetigo contagiosa—but the primary cause of their trouble was seborrhoeic. Amongst other things he pointed out that a characteristic sign was a chronic, often crusted, sore on the sulcus half-way between the lower lip and the point of the chin. One cannot cure seborrhoeic dermatitis with penicillin, and these 'seborrhoeic impetigos' remained a problem until the end of the war.

These, then, are but two examples of the wide variety of often unrecognized seborrhoeic maladies, and most dermatologists can give examples of the seborrhoeic state as the 'onlie begetter' of a variety of cutaneous lesions which cause a great amount of human suffering and sick wastage.

It seemed therefore that if at St. Bartholomew's we could discover some basic facts about sebum and its secretion many things would be yielded unto us.

We had little to go on, although in 1936 H. W. Barber in a summary of his views concerning the seborrhoeic state had defined seborrhoea as 'an excessive and altered secretion from the sebaceous glands and (i.e. associated with) a change of the composition of the fat in the horny layer (of the epidermis)'. At that time it was believed that seborrhoea was the most important factor leading to seborrhoeic eruptions. The bacteriology of seborrhoeic eruptions had been a cause of controversy for two generations, and we decided, at first at any rate, to leave this matter alone; perhaps we were wise for I. Martin-Scott (1952) has since shown that so far as *Pityrosporon ovale* is concerned it is probably a non-pathogenic saprophyte which can be cultured

now believe that the vitamin is not formed in the skin, but is photo-synthesized on the surface of the epidermis from a precursor (probably 7-dehydro-cholesterol); oddly enough this sterol has not yet been demonstrated in sebum. Once formed the vitamin is rapidly absorbed—though how this happens is not known—and probably reaches the blood stream in the corium.

SEBUM

At St. Bartholomew's Hospital and Medical College we have, in recent years, taken much interest in sebum. Professor A. Wormall and myself agreed that we would endeavour to stimulate research under our joint aegis, and we have been fortunate in securing the generous interest of the College in our efforts. Our plan was that I should suggest the problems from the clinical side and would discuss them with Wormall who would assess them from the biochemical standpoint. In V. R. Wheatley we have found a very competent junior colleague who first studied the clinical problems with me and then, under Wormall's direction, bent to the arduous task of solving them in the laboratory. A little later I. S. Hodgson-Jones was able to add his contribution to the project. The first problem was dictated by the enigma of seborrhoeic dermatitis. This malady is common in most—probably all—parts of the world. Its manifestations are so protean that I know of no adequate comprehensive survey of them in the literature. To give but two examples. When P. F. Borrie and I worked together at what was then called the Wellhouse Hospital, we noted on many occasions a chronic eczematous eruption on the lower legs of elderly men which to the casual eye was just a variety of hypostatic eczema. These patients, however, were seborrhoeic subjects, often with chronic petalloid seborrhoeic dermatitis on their chests, dandruff in the scalp, and with the stigmata of the seborrhoeic state which H. W. Barber has so excellently described. In our view these rashes on the legs—on limbs often tightly ensheathed in woollen undergarments and socks—were but manifestations of seborrhoeic dermatitis. The second example of the nature of seborrhoeic dermatitis was observed by J. H. Twiston Davies during World War II. The incidence of eruptions recorded as

secondly it contains 30 per cent of free fatty acids—an unusual feature in a naturally occurring fat; thirdly, for the first time the unsaturated hydrocarbon squalene was shown to be present in significant amounts in a normal human tissue fat; fourthly the presence of a large proportion of paraffins was surprising as these rarely occur in animal tissues. In addition we failed to discover vitamins A, B, C, or K, pro-vitamins D2 and D3, or B-carotene, and were able to detect only a trace of vitamin E· we had expected that fat-soluble vitamins or pro-vitamins would be present.

The discovery of squalene has aroused some interest: not only did many industrial chemists working for cosmetic firms write for reprints of our report—so that we expected to see 'Squalene Face Cream' being marketed in New Bond Street and Fifth Avenue—but it also stimulated inquiry whether squalene is a precursor of cholesterol; we note however that certain laboratory animals do not have squalene in their sebum although they appear to secrete cholesterol from their sebaceous glands.

TABLE 2. *Comparison of the Composition of Human and Sheep Sebum*

	Man	Sheep
Free fatty acids	30%	trace
Triglycerides	+	–
Waxes	+	+
Cholesterol and its esters	+	+
Lanosteryl esters	–	+
Squalene	+	–
Other hydrocarbons	+	trace
Nature of free and combined fatty acids	*n*-series (odd and even)	*n*-series (even) Branched-chain (iso- and anteiso-)

In an attempt to discover the origin of squalene, Wheatley (1953) analysed sebum obtained from rats, rabbits and guinea-pigs; he found that unlike that derived from humans, this sebum was not a mixture of fats and waxes, for only a trace of triglyceride could be detected. This observation was of some interest because it has for long been known that sebum from the sheep is also a wax and not a fat (in parenthesis, clinicians will note

from normal adults as frequently as from seborrhoeic subjects and is not usually found alive in cases of acute seborrhoeic dermatitis.

We noted that there was some controversy concerning the mechanism of sebum secretion. In 1934 J. Boeke had described a nerve plexus investing sebaceous glands: he thought it was derived from the sympathetic system, but noted that the individual cells of the glands were not innervated. G. R. Cameron and W. G. Spector (1953) have recently reached the conclusion that whilst experiments in animals afford some evidence in favour of nervous control, on the whole there is little in favour of the view in the case of man; our work has done something to support this opinion—but I still find it very difficult to believe that these glands are independent of nervous control, and at the outset I had hoped we would find evidence in favour of Boeke's view.

Thus it was some four years ago we embarked on our project. Wheatley's first task was to investigate the chemical composition of human sebum. The material was obtained by immersing the forearms of volunteers in acetone and recovering the sebum thus extracted from the liquid. The material he obtained was an amber-coloured wax-like substance resembling the ear wax of Europeans. Analysis showed that its average composition was as shown in the following table.

TABLE 1. Average Composition of Human Sebum (MacKenna, Wheatley and Wormall, 1950; 1952)

	Per cent
Free fatty acids: unsaturated	15
saturated	15
Triglycerides	32 5
Waxes (including cholesteryl esters)	15
Sterols: Cholesterol (free)	2·5
Cholesterol (combined)	(2 5)[1]
Other sterols	2 5
Squalene	5
Paraffins	7 5
Unidentified (including oxidized squalene)	5

[1] Included in waxes

There are several important features about this analysis. First, human sebum is shown to be a mixture of fats and waxes;

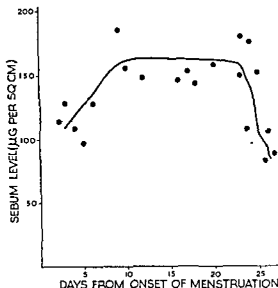


FIG 5. Variations in sebum levels during the menstrual cycle Measurements made on the upper back of a woman aged 18 years.

Finally an attempt was made to show changes in the sebaceous activity in seborrhoeic dermatitis. The sebum levels of certain areas were estimated, and the sebum analysed. The levels in subjects suffering from seborrhoeic dermatitis were very little different from those in normal subjects but the results showed a slight but definite alteration in the composition of sebum in cases of this malady (Table 4).

TABLE 4. Some Chemical Determinations on Sebum

	Normal	Seb. Derm.
Acid No.	44 ± 15	44 ± 16
Iodine No	67 ± 7	59 ± 11
Cholesterol (%)	4 1 ± 1·3	4 6 ± 2·5
Squalene (%)	8 5 ± 1·7	7 2 ± 2 3

More detailed investigation of the chemical changes in sebhorrhoeic dermatitis and other skin diseases is now required;

that although Lanolin is, in practice, a satisfactory substitute for human sebum, its chemical composition is surprisingly different)—but sheep's sebum differs considerably in composition from the sebaceous secretion of rabbits, rats and guinea-pigs. We learned from these investigations that for our purposes we would have to concentrate on sebum derived from human beings; we also had to note certain problems: for example Flesch and Goldstone (1952) showed that when squalene was applied once to the skin of rabbits and guinea-pigs loss of hair occurred some 10 days later. Within a few weeks the fur is replaced. Unfortunately we found that squalene does not similarly cause epilation in human subjects, but this was to be anticipated, squalene being a normal constituent of human skin fat.

Besides studying the nature of sebum it was necessary also to study the quantitative activity of the sebaceous glands. For such an investigation the amount of sebum covering a measured area of skin has to be estimated—this is termed the sebum level. Wide individual variations are found in the sebum levels of different areas of the body. There is, however, a definite gradation in the sebum levels of different areas, the highest occurring in the scalp and the lowest in the extremities (Table 3).

TABLE 3. Some Sebum Levels of Various Sites of the Body (Hodgson-Jones and Wheatley, 1952)

Site of Body	Average sebum level (μg /sq. cm.)
Forehead	212 ± 73
Chest	120 ± 61
Back	106 ± 56
Axilla	84 ± 59
Groin	75 ± 28
Abdomen	67 ± 45
Arm	58 ± 34
Leg	36 ± 19

It was observed that in men the sebum level for a given area remained fairly constant from day to day, but in women the sebum level follows a recurrent pattern during the menstrual cycle (Figure 5).

organisms; the finding has caused considerable surprise, which has been increased by the report that the acne bacillus, which is an anaerobe, is numerically the preponderant resident organism of the skin.

In the matter of relative inviolability of the skin to transient microbic infection we have much of practical experience, some knowledge, and a good deal of theory. Let us for a moment consider two forms of infection of the skin, usually attributed to staphylococcal infection, namely furunculosis and impetigo.

Boils are—or used to be—occupational maladies of house physicians and house surgeons towards the end of their period in office. They occur readily in persons who have had a good deal of emotional stress, too little fresh air and exercise, and who have lived on a diet too rich in carbohydrates. We can, in our ignorance, omit considering the factor of stress; but I suggest quite dogmatically that boils seldom occur in healthy subjects with reasonably sun-tanned skin living an open-air life: that is in those persons whose environment is conducive to adequate cutaneous metabolic activity. The dietetic factor is important. In cases of chronic furunculosis in non-diabetic subjects, the patient often takes an enormous amount of carbohydrate; many such persons will admit having an average of 8–10 cups of tea a day with 3 teaspoonfuls of sugar in each cup: much of their solid food is also excessively sweetened; their lesions often disappear when they submit to a reasonable curtailment of their consumption of sugar and starch. On clinical grounds the theories of E. Urbach and J. W. Lentz (1945) deserve consideration, but I am informed that no one has been able to confirm or disprove their views. These workers, particularly Urbach, found that the average sugar level in the skin was 58 mgm. per 100 grammes and regarded a level above 68 mgm. as pathological They suggested that there is an independent, intermediary carbohydrate metabolism in the integument and used the term independent cutaneous glycohistechia to designate a rise in skin-sugar level not accompanied by or preceded by a rise in blood-sugar level; they considered that this occurrence is not related to the general sugar metabolism of the body and therefore is not related to ordinary diabetes mellitus.

this will entail the development of 'ultra-micro-chemical-methods' of analysis in order that the composition of sebum from small areas of skin can be examined in detail. In view of the dietary restrictions which some clinicians customarily impose on patients suffering from seborrhoeic disorders—and for other reasons—it will be interesting to investigate the effect of diet on the composition of sebum. Also there are the matters of the origin of squalene, its function, vitamin D formation and other problems which Wheatley hopes to investigate in the near future.

Infection of the Skin

The penultimate subject to which I wish to draw attention is a consideration of the bacterial flora of the skin and some views concerning the protection of the integument from bacterial invasion. Clinically this matter is not so remote from the problems of seborrhoea as it may appear to be to the scientists in my audience, for there is some clinical evidence to show that in the seborrhoeic state there is frequently some degree of breakdown in the mechanisms of self-disinfection of the skin.

D. M. Pillsbury and A. M. Kligman (1953) have recently reviewed the topic. They point out that the skin is never sterile, and cannot be rendered so by any practical means; a variety of nitrogenous and lipid constituents enable it to support a luxuriant flora, which is predominantly bacterial, on its surface and in the orifices of its specialized glands.

They believe that individual variability of the cutaneous flora of different persons is perhaps a characteristic feature, for there are real differences in the average number of organisms supported by the skin of different individuals. Certain persons may unquestionably be characterized by having consistently high counts whilst others—just as consistently—have low counts. There seems to be some agreement amongst those who have studied these matters that most bacteria are not on the skin surface, in or between the living cells of the epidermis, but lie within the openings of the glandular appendages. This may explain why Evans and his colleagues (1950) have been able to show that the anaerobic bacteria on—or in—the unspecialized glabrous skin are greatly in excess of the aerobic

SYMPATHY: PSYCHOLOGY

At the beginning of this lecture I suggested that the scientific approach to dermatology necessitated, among other things, a study of how the skin reacts to morbid agents in disease. In connection with this, I would like to suggest that a study of 'sympathy' and homology might be of some interest.

For many years ophthalmic surgeons referred to sympathetic ophthalmia meaning an inflammation of the uveal tract of one eye, due to the effects of a similar inflammation in the other; it is not very common nowadays, but is still recognized as a clinical entity. The suggestion I wish to make is that in dermatology there is a sympathy or homology between some areas of skin, so that when the first area is affected, another definite area is likely to develop some form of secondary eruption. I propose using the word 'sympathy' when referring to interrelated phenomena occurring on analogous areas of the skin, and homology in reference to interrelated phenomena occurring on differing areas of the integument.

There is, I submit, a small amount of clinical evidence to suggest that 'sympathy' may properly be regarded as a dermatological phenomenom as well as an ophthalmic one. The subject is difficult, because it is linked with the problem of distribution of skin eruptions (of which we know nothing except the visible facts) and the psychological matter of unilateralness and bilateralness

If a man develops tinea pedis, he may later develop an 'id' eruption, in which event the 'id' is more likely to be on the hands than on any other area. In my view that is one example of homology. The matter may be lightly dismissed by saying that both hands and feet are peripheral areas remote from the heart. If he has a sluggish circulation in his hands the toxic products from his feet will naturally concentrate in the vessels of his hands and thence damage the skin of his hands; but I would suggest that this argument is not entirely valid, for if he develops tinea cruris without tinea pedis, although the tinea cruris would be caused by one of the same varieties of fungus that attack the feet, it is most unlikely that he would develop an 'id' eruption.

I am well aware that these views have been adversely criticized and even ridiculed; from the solely clinical point of view they provide an explanation of a phenomenon which one has observed in many patients; and if Urbach and Lentz were wrong, it would be interesting to know the correct explanation of the diatetic factor in furunculosis to which I have referred.

From the foregoing observations, we can assume that stress, diet, environment and the seborrhoeic state are factors which the clinician has to remember when he considers cases of the two named infections.

Now there are two principal theories concerning the 'degerming' of the skin (I hasten to add that 'degerming' is an accepted nonce-word of technical jargon). The first is a chemical matter and refers to the acid mantle of the skin: Marchionini (1928 and 1929) thought that acids in sweat and the low isoelectric point of keratin were of much importance. Burtenshaw (1948) cautiously suggested that long-chain fatty acids were chiefly responsible. The second theory is physical: Norton and Novy (1931), Rebell *et alia* (1950) have favoured the belief that the self-disinfecting mechanism is due to desiccation. Ricketts, Squire and Topley (1951) have reconciled some of the divergent points of view by showing that desiccation and chemical sterilization may both be important, but the importance varies in accordance with the organism that is being studied. Thus *E. coli* and *Pseudomonas* are removed easily from the skin by desiccation but multiply in moist areas. Whilst there are many gaps in our knowledge Pillsbury and Kligman (1953) have suggested that desiccation and not chemical sterilization would appear to be a significant factor in causing the disappearance of Gram-negative organisms: both desiccation and chemical effects influence the rate of disappearance of *Staph. aureus*.

Time will show whether the factors I have thus briefly mentioned are the sole ones which we have to consider in regard to self-disinfection of the skin, but it is worth while emphasizing that the mechanisms are extremely efficient, as witness the great difficulty which Sheehan and Fergusson (1943), Bigger and Hodgson (1943), and others have had in their efforts to produce impetigo contagiosa experimentally on human skin.

Now, for the present argument, I suggest that the pattern of skin eruption may be influenced by sympathy and homology to a degree which we do not realize. They may at any rate be invoked to explain the symmetricity and pattern of certain eruptions.

But, it may be said, the symmetry in distribution of a secondary syphilitic rash or psoriasis, and the pattern of, say, chronic lupus erythematosus are more features of the disease process *itself than anything to do with the patient. For present purposes* I suggest that is too sweeping a statement, and would like to elaborate this view.

When *Staphylococcus pyogenes* is on a culture plate, it may cause a few changes in the medium, but it does not cause a boil. It will only form a boil if it achieves a suitable breeding-ground in a pilo-sebaceous follicle and if the host reacts to the stimulus of the microbe. The reaction to the daemon is the concern of the host and it is this which produces the boil. If the host does not respond, commensalism may ensue, or the daemon may be destroyed by those processes of self-disinfection which we have been discussing. It is probably only given to a few daemons—the bacilli of anthrax and glanders for example, granted that they are really virulent—almost invariably to produce the signs and symptoms with which they are accredited.

Now microbic infection is the easiest example by which I can emphasize the importance of the host's reaction in producing what we call disease.

We tend to postulate that some daemon or other—and the daemon may not necessarily be bacterial or viral, but may be a metabolic disorder, an allergen, a hapten, or even an involutionary process—some daemon or other is responsible for the host reacting to produce psoriasis, lichen planus, pemphigus or dermatitis herpetiformis. But these diseases can be described in text-books as following certain well-defined trends and having certain well-defined patterns of eruption, not so much because the daemon causing say pemphigus vegetans is, we presume, the same in all cases, but because the reactions of the hosts tend to be much the same in most cases. If this is accepted it means that the paragraphs concerning aetiology in our text-books will

There is therefore, I think, some slight evidence to suggest that there may be some degree of homology between the feet and the hands, so that if the former sustain a fungus infection the latter are likely to react.

When suspender dermatitis occurs on the thighs (the eruption being caused by an allergic hyper-sensitivity to the metallic plating of the suspenders), it is not unusual to find an itching, papular, secondary eruption on the surfaces of the upper limbs adjacent to the elbow. Why should this be the site of election for the secondary eruption?

Plate VIII, Figure 6 and Plate IX, Figure 7 show cases of a fairly acute eczema limited to the limbs, and of lichen planus. The latter illustrates 'sympathy', the former both 'sympathy' and 'homology'.

Besides the homology which seems to exist between the lower and upper limbs, there is also some degree of sympathy between one leg and the other. Take a man with an oozing dermatitis on his right leg: seal in the rash with soft paraffin and it is quite probable that he will get a secondary rash on the other leg. Similarly there may be sympathy between one arm and the other, one foot and the other, one hand and the other. I am not certain that the homology that seems to exist between the feet and the hands, and the legs and the arms occurs as easily in reverse; for, if a patient develops an eruption on the upper limbs it does not seem to be so likely that he will get an homologous eruption on the lower ones. If the conception of sympathy and homology is accepted as valid, then it may be that in certain maladies such as generalized exfoliative dermatitis of the Pautrier-Worringer type, for example, sympathy and homology play some part in perpetuating the eruption, and something might be done to exploit these factors in treatment. But we know little or nothing about the mechanisms which are involved, and therefore are powerless to do much except theorize at present Nevertheless, out of theory, investigation may ensue: it is part of the scientific approach—just as it is part of that approach very carefully to scrutinize any such ideas without bias, to see if the clinical observations on which they are founded are valid.

daemon, when his cutaneous resistance breaks down he develops a recurrence of say his lichen planus, his endogenous eczema, or his psoriasis. The type of the eruption will depend both on the nature of the daemon and the resistance and reaction of the host; the pattern of the eruption and its extent may be influenced to some degree by the factors of homology and sympathy, as well as by the extent of breakdown of resistance.

How—in dermatology—is there a breakdown in resistance? We have considered breakdown in relation to bacterial infection. In eczema, in psoriasis, in lichen planus, the breakdown may be due to somatic causes. It may also be due to those factors of stress to which Selye has devoted so much attention. And here it is that we approach the influences of the mind on

this aspect which was broached to a surprising degree by the Victorians and scorned by a later generation, slowly become accepted by the majority. Much of the present difficulty is that the *scientific* discipline is one whereby the individual is taught to assess his results not by the light of his own experience but by comparison with controls and standards which are acceptable to and may have been set by others.

Controls and standards are not so readily available to the psychiatrist; his method is perforce 'individualistic'; his data cannot be measured; possibly the modern scientist who finds himself, in physics at any rate, dealing with intangibilities, is for this reason more tolerant of the psychologist than he used to be. But although each submits to different disciplines, the place of the dermatologist is between the two: he can learn from both—his knowledge of his Art must teach him when to request help from one and when from the other. But eventually it will be the triumph of the anatomist, the biochemist and the physiologist to explain the exact mechanisms whereby the skin can—in certain individuals—be the mirror of the emotions.

eventually acquire a different 'slant'. Impetigo contagiosa will not, for example, be defined as a disease due to invasion of the epidermis by *Staphylococcus pyogenes*, but as a cutaneous reaction provoked by the organism in certain subjects whose mechanism of cutaneous self-disinfection has broken down in such and such a way, because of such and such factors, but whose follicular resistance is sufficiently strong, for such and such reasons, to prevent them from having boils or carbuncles instead of impetigo.

Disseminated lupus erythematosus is usually associated with a severe generalized cutaneous eruption. At the Mayo Clinic they are sure that many women having malaise, arthralgia and arthritis without rontgen demonstrable bony changes, albuminuria, leucopenia, a high E.S.R. and reversal of the albumen globulin ratio but with no cutaneous eruption, are suffering from this malady. These women have no dermatological signs probably because their cutaneous reactions are different from most of their fellow sufferers.

Psoriasis is a malady in which the pattern of inheritance has baffled the geneticists. I suggest that all persons in a psoriasis family have the daemon which they pass on to their offspring; but that—in most families—only a minority of these offspring react with cutaneous lesions. They may react in other ways and it would be interesting to see if, for example, there is a pattern in the medical histories of so-called non-psoriatic members of these families: are they prone, for example, to leukoplakia or is the daemon able to change its attack so that instead of concentrating on ectodermal tissue it can eventually wreak its vengeance on structures or organs derived from the mesoderm or endoderm? Are there, in fact, unsuspected relationships between some cutaneous diseases occurring in one generation and visceral, endocrine, cardiovascular and skeleto-muscular diseases occurring in future generations? If this is not just wild speculation, then the phenomena occurring in atopic families (that is the asthma, eczema, hay-fever, migraine groups) are not so unique as we suppose.

We began this part of the lecture by considering problems of homology and sympathy. Granted that a man possesses a

WEDDELL, G. (1953). *Modern Trends in Dermatology* Second Series. Butterworth, London.
WEDDELL, G. and PALLIE, W. (1954). *Quart J. misc. Sci.* **95**, 389.
WHEATLEY, V. R. (1953) *Biochem. J.* **53**, xxi.
WITTKOWER, E. and BRIAN RUSSELL (1953). *Emotional Factors in Skin Disease.* Paul B. Hoeber, New York.
WOOLLARD, H. H., WEDDELL, G. and HARPMAN, J. A. (1940). *J. Anat., Lond.*, **74**, 413.

REFERENCES

BARBER, H. W. (1936). *Taylor's Practice of Medicine*. 15th edition. J. & A. Churchill, London.

BIGGER, J. W. and HODGSON, G. A. (1943). *Lancet*, **i**, 544.

BOEKE, J. (1934). *Z. mikr.-anat. Forsch.* **35**, 551.

BURTENSHAW, J. M. L. (1948). *Modern Trends in Dermatology*. First Series. Butterworth, London.

CAMERON, G. R. and SPECTOR, W. G. (1953). *Ibid.* Second Series. Butterworth, London.

DOWLING, G. B. (1953). *Ibid.* Second Series.

EVANS, C. A., SMITH, W. M., JOHNSTON, E. A. and GIBLETT, E. R. (1950). *J. invest. Derm.* **15**, 305.

FLESCH, P. and GOLDSTONE, S. B. (1952). *J. invest. Derm.* **18**, 267.

HODGSON-JONES, I. S., MACKENNA, R. M. B and WHEATLEY, V. R. (1953). *Brit. J. Derm.* **65**, 246.

HODGSON-JONES, I. S. and WHEATLEY, V. R. (1952). *Biochem. J.* **52**, 460.

LEACH, E. H. (1952). *Brit. J. Derm.* **64**, 183

LEIDER, M. (1949). *J. invest. Derm.* **12**, 187.

LEWIS, T. (1942). *Clin. Sci.* **4**, 365.

MACALPINE, I. (1953). *Modern Trends in Dermatology*. Second Series. Butterworth, London.

MACKENNA, R. M. B., WHEATLEY, V. R. and WORMALL, A. (1950). *J. invest. Derm.* **15**, 33.

MACKENNA, R. M. B., WHEATLEY, V. R. and WORMALL, A. (1952). *Biochem. J.* **52**, 161.

MARCHIONINI, A. (1928). *Schweiz. med. Wschr.* **9**, 1055

MARCHIONINI, A. (1929). *Arch. Derm. Syph., Wien*, **158**, 290.

MARTIN-SCOTT, I. (1952). *Brit. J. Derm.* **64**, 257

MUCHOW, H. G. K. (1925). *Tabulae Biologicae*. Edited by W. Junk. **2**, 469.

NORTON, J. F. and NOVY, M. F. (1931). *Amer. J. publ. Hlth.* **21**, 1117.

PILLSBURY, D. M. and KLIGMAN, A. M. (1953). Modern Trends in Dermatology. Second Series. Butterworth, London.

REBELL, G. C., SAINT PHALLE, M. and GINSBURG, D. (1950). *J. invest. Derm.* **14**, 247.

RICKETTS, C. R., SQUIRE, J. R. and TOPLEY, E. (1951). *Clin. Sci.* **10**, 89.

ROTHMAN, S. (1941). *Physiol. Rev.* **21**, 357.

SEQUEIRA, J. H., INGRAM, J. T. and BRAIN, R. T. (1947). Diseases of the Skin. Fifth edition. J. & A. Churchill, London

SHEEHAN, H. L. and FERGUSSON, A. G. (1943). *Lancet*, **i**, 547.

TWISTON DAVIES, J. H., DIXON, K. and STUART-HARRIS, C. H. (1945). *Quart. J. Med.* **56**, 183

URBACH, E. and LENTZ, J. W. (1945). *Arch. Derm. Syph., Chicago*, **52**, 301.

directed towards a search for regularities and concurrences in behaviour and aims at establishing principles which describe them and from which predictions can be made.

Hypotheses

Experimental studies of psychopathological states began during the early years of this century as a result of observations made in Pavlov's laboratory. Here investigators, quite incidentally to the main purposes of their research, produced behaviour abnormalities which, in their opinion, were analogous to the neuroses in man (Pavlov, 1928). The characteristics of the conditioned salivary response in dogs were being studied. One animal had been conditioned to salivate when a circle of light was projected on a screen in front of it and to inhibit salivation to an ellipse of the same size and intensity. This differential conditioned response continued to be evidenced as the shape of the ellipse was brought closer and closer to that of the circle until a point was reached when the ellipse was nearly circular and dramatic changes occurred in the animal's behaviour. As Pavlov (Pavlov, 1928) described the event, the differential conditioned response '. . . not only disappeared spontaneously, but caused the loss of all earlier differentiations . . . the dog which formerly stood quietly on his bench, now was constantly struggling and howling'. Breakdown of the animal's usual behaviour had been produced in a laboratory under conditions which investigators could control. Here, as has been so frequently the case in the history of science, more or less casual observations opened up a new field of experimental investigation.

As is characteristic of early stages of development in all sciences, these and similar observations gave rise to a variety of hunches, in this case hunches regarding the conditions essential to the production of behaviour abnormalities and the characteristics of the abnormalities themselves. Pavlov couched his hunches in physiological terms, in terms of processes having their locus in the cerebral cortex. Psychologists emphasized relations between behaviour disorders and certain critical environmental and organic conditions. All found it necessary to refine their hunches by systematic preliminary studies before

XXI

Experimental Psychopathology

ROGER W. RUSSELL

Introduction

I INTEND to limit the discussion which follows to one, rather young, branch of experimental psychology, consideration of which is certain to illustrate the kinds of problems faced by those with an experimental bent who study the behaviour of living organisms. As I understand my present obligation to you it is to indicate how the methods of experimental science have been applied towards the solution of these problems. The examples I have chosen to illustrate my discussion all deal with some aspect of behaviour or its neurophysiological correlates. In a number of cases they are the contributions of persons subscribing to disciplines other than psychology—physiology, zoology, and biochemistry to name three. Psychologists have no monopoly of interest in the behaviour of living organisms. They may differ from other interested parties in those aspects of behaviour which constitute their primary concern and in the terms in which their theories are formulated, but all have in common the rigorous requirements of the experimental methods of science which determine their approach to research.

The researcher who employs experimental methods in his study of behaviour starts, as do other scientists, with certain articles of faith Among these basic assumptions are two of particular concern to us at this moment. First, what happens in nature—and this includes the behaviour of living organisms—occurs in accordance with principles or laws. Second, these principles or laws are such that human reason can discover them. He who starts with these attitudes finds his attention

who, a number of years ago, wrote in his *General Introduction to Psychoanalysis* (Freud, 1920): '. . . As to the neurotic symptoms, we already know that they are the result of conflict.' The evidence supporting statements of this kind has come mainly from *ex post facto* reasoning regarding conditions which may have affected the development of behaviour disorders already present. The methods of experimental psychopathology make it possible to begin observations on subjects who do not show symptoms of abnormal behaviour, to expose these subjects to known and clearly defined conflict situations, to measure changes in behaviour elicited by such exposure, and to do all this under conditions over which the experimenter has complete control. The real value of the experimental approach lies in the fact that these advantages allow more precise analysis of conditions critical to the production of behaviour disorders, thus increasing the confidence which may be placed in any general principles that may arise from the observations made.

A few examples may help to illustrate how hypotheses regarding the aetiology of behaviour disorders have been put to experimental test. Earlier I pointed out how analysis of the Pavlovian techniques for producing 'experimental neurosis' in animals suggests that conflict may appear in its simplest form as competition between two incompatible response tendencies elicited simultaneously. One series of experiments (Miller, 1944), using rats as subjects, was designed to investigate the validity of this hypothesis and to identify those attributes of the environment and of the organism which are critical to the disorganization of behaviour. Hungry animals were first trained to run from one end of a simple straightaway to the other in order to reach food. When this approach response was well established the animals were given an electric shock on reaching the food-end of the apparatus. As would be expected, the shock produced an avoidance response, and a simple conflict, which we might refer to as an approach-avoidance competition, resulted. The investigators had predicted on the basis of preliminary studies that behaviour under these conditions would be characterized by vacillation round a point in the straightaway where the competing tendencies were of approximately equal strength,

proceeding to final experimental verifications. For example, one way of analysing the Pavlovian observations may take the following form. The differential conditioned responses originally established had two major components. Presentation of the circle, always accompanied by food, led to the development of a conditioned response in which the circle became a signal for salivation. On the other hand, the ellipse, which was never reinforced by food, became a signal for the inhibition of salivation. In other words, the two stimuli, through training, came to elicit responses which were completely incompatible. As the ellipse was made more circular a point was finally reached when the dog could not discriminate between the two shapes. At this point the stimuli simultaneously elicited two incompatible responses and it was under these conditions that the breakdown in normal behaviour occurred. Following this hunch further several investigators studied the possibility that the essential condition in the production of behaviour disorders is conflict between two or more mutually exclusive responses. The investigators found it necessary to make certain basic assumptions regarding the characteristics of the competing response tendencies and to test these assumptions by preliminary observations. Such refinement of hunches eventually led to the statement of hypotheses which could be put to experimental test. The clear statement of hypotheses is an essential preliminary step in any experimental study. Darwin once wrote: 'How odd it is that anyone should not see that all observation must be for or against some view, if it is to be of any service.' It is not infrequent in psychology to find those who refuse to recognize this important point, who declare that they are interested in 'facts' and 'facts' alone. Huxley said to those in another science subscribing to such an attitude: 'Those who refuse to go beyond fact rarely get as far as fact. . . . Almost every great step (in the history of science) has been made by the "anticipation of nature", that is, by the invention of hypotheses which, though verifiable, often had very little foundation to start with.'

The functions of conflict in the aetiology of human behaviour disorders have been given central roles in the writings of a number of prominent clinical psychopathologists including Freud,

Not all analyses of situations in which behaviour disorders have appeared lead to the conclusion that some form of conflict is involved. Let me introduce the point I am about to make by describing an experiment. In this experiment (Barker, Dembo, and Lewin, 1941) thirty children, ranging in age from two to six years, served as subjects. They were observed individually on two different occasions. The first observation took place in a playroom under conditions involving unrestricted play. The second experimental period took place in the same room, but under conditions in which certain attractive toys were visible but not always accessible. During this second period the subjects were first allowed free play with all toys, then separated from the toys by a wire partition, and, finally, again given opportunity for free play. During both periods the children's behaviour was recorded in detail. The effects on behaviour of imposing a physical barrier between the subjects and the toys to which they had previously had free access were most striking Marked regressions in constructiveness of play, equivalent in some cases to twenty-four months of mental age, were observed. Changes in emotional expression accompanied this regression and were evid

crea

aggr

unwarranted extent could this experimental procedure be placed in the same category as those I have previously described. The main feature of this experiment is that disorganization of behaviour occurred under conditions in which the activities of the subjects were restricted by an attribute of the environmental situation which they did not have the capacity to surmount. Conditions such as these have been referred to in the clinical and experimental literature as 'frustration' From a large number of possible examples I shall select only one more to illustrate how the basic conditions of frustration have been subjected to experimental analysis. A very extensive series of experiments has been concerned with the effects of frustration on the development of 'abnormal fixations' (Maier, 1949). The subjects, in this case rats, were forced to jump from a small platform on which they were confined against one of two doors in a

a prediction which in general was confirmed by the experimental results. Behaviour in this conflict situation was non-adjustive in the sense that it led neither to satisfaction of the tissue conditions underlying the hunger drive nor to escape from circumstances potentially injurious to the animal. Other experiments (Gantt, 1944; Masserman, 1943) in which animal subjects have been exposed to conflicts of this simple approach-avoidance type have produced even more striking behaviour disorders, including 'phobic' responses, 'regressions', behaviour rigidity, and extensive neurophysiological changes. Still other research has shown that approach-avoidance is only one of three basic types of conflict situation and that disorganization of behaviour may also occur when the competition involves either two avoidance responses or, under somewhat more complex circumstances, two sets of approach-avoidance responses.

The fundamental part which these conditions play in disorganizing behaviour has been demonstrated by experiments on human as well as animal subjects, although the conflicts to which the human subjects were exposed have of necessity been much more mildly stressful. For example, one investigator (Barker, 1942), using young boys as subjects, instructed them to move a lever towards the name of the one of two liquids they would prefer, and, in some cases, were required to drink. The solutions, ranging from orange juice to saturated salt solution, had been ranked in order of desirability by each subject prior to the experiment proper. The results indicated that there existed '... an inverse relationship between the time required to resolve a conflict and the "distance" between the alternatives in the preference series'. A similar relation existed between the amount of response vacillation and 'distance' in the series. Conflicts between undesirable alternatives—the double-avoidance situation—produced greater disorganization of behaviour than conflicts between desirable alternatives. The disorganization was temporary and limited, but it did [illegible]

[illegible]

[illegible] support of the basic principles I have been describing.

Independent Variables

The hypotheses I have been discussing have a certain similarity of form. They all consist of statements regarding relations between two types of variables, referred to in experimental science as 'independent' and 'dependent' variables. As the hypotheses were stated these variables were related as 'cause' to 'effect'. One of the most important requirements of the experimental method in science is that the investigator must have under his control the independent variables he is studying and that he be able to manipulate them systematically. In contrast the non-experimental methods of science are applicable in situations where the researcher is not able to manipulate independent variables but must search for instances in nature where systematic variations occur spontaneously. Unfortunately this differentiation is not always adhered to; the term 'experimental' is loosely applied to observations of almost any variety and thus loses restrictions which, when accurately applied, give experimental methods their distinctive character.

Where control and manipulation of independent variables are possible careful research design can provide considerable advantages for the researcher who seeks to uncover the dynamics of 'abnormal' behaviour in the most rigorous manner. He can produce the behaviour at a time and place which make it possible for him to be fully prepared for accurate observation. He is able to watch the behaviour as it develops and thus is not limited by dependence upon *ex post facto* reasoning regarding events which may have influenced the development. If his research is adequately planned he can repeat his observations and report conditions in such a way as to allow himself or other observers to reproduce the research. The desirable scepticism with which scientists view each other's work makes this latter point a necessity. Perhaps the greatest advantage of all is that the experimental psychopathologist can control the conditions which he believes may affect the behaviour he is studying and can systematically manipulate them.

Previous examples have shown how certain attributes of the environment and of the organism, when manipulated in such a manner as to interfere with the organism's adjustment, produce

panel before them. Under ordinary learning conditions a jump against the correct and unlocked door, marked in a particular manner, led to reward. Jumping to the incorrect and locked door was punished by bumping against the panel, falling into a net slung below, and receiving no reward. When the problem was made insoluble by locking and unlocking doors in a random manner the majority of subjects developed fixations which usually took the form of responses to one particular door. These fixations were strikingly rigid when attempts were made to break them. Subjects showing such fixated behaviour frequently did not alter their responses sufficiently to adjust to new and soluble problems, even after several hundred trials. During these trials the animals gave conclusive evidence of having 'learned' the soluble problem in that they would approach the correct door on each trial. If the correct door were on the fixated side they would jump through it without hesitation, but, if the correct door were on the non-fixated side, they would, after long response latencies, jump to the incorrect door despite the punishment which inevitably followed. This was quite a different pattern of responses from that of subjects learning the same behaviour under conditions of no frustration, who in a short time altered their behaviour to conform with the requirements of new problems when they were presented.

When the relevant experimental studies are analysed (Russell, 1953) it seems apparent that behaviour disorders are produced under conditions in which some event interferes with the adjustment of an organism to environmental or organic changes. Adjustive behaviour appears to be obstructed under either of two general sets of critical conditions. Under one the nature of the obstruction is such that the organism does not possess the capacities or skills to surmount it. Such conditions are usually referred to by the generic term 'frustration'. Under the second, competition arises between two or more incompatible response tendencies resulting in delay or failure in making a compensatory response, conditions generally referred to as 'conflict'.

individual differences in stress tolerance that they have become a central problem in experimental psychopathology. Attempts to understand the nature of these differences have searched for contributing conditions either in the organism's life history prior to exposure to stress or in its hereditary background.

There is at present some experimental evidence from animal research that events occurring in the individual's earlier life history have significant effects on his susceptibility to breakdown and on the nature of that breakdown once it does occur. There is evidence (Hunt, 1941) to indicate that rats subjected to feeding frustration during infancy develop abnormal food hoarding behaviour in later adult life. Artificially nourished puppies and calves which are not permitted to indulge in nipple-sucking to the point of satiation develop abnormal behaviour patterns in the form of licking their paws, other dogs, or inanimate objects to a degree dangerous to the animal's health, behaviour not evidenced by control animals not so frustrated (Levy, 1934).

Approaches to the question whether or not hereditary factors contribute to individual differences in the development of behaviour disorders have been made by application of the methods of experimental genetics in which hereditary and environmental factors are controlled and systematically varied. For instance, selective breeding experiments have suggested that hereditary factors do play an important part. One study (Hall, 1938), using rats as subjects, exposed the initial sample of animals to a mildly stressful situation in which their behaviour could be observed and measured. The animals at both ends of the distributions of these measures were then inbred and their offspring tested in the same manner. This procedure was continued until by the eighth generation two groups had been isolated with very little overlap between them, one very susceptible and one very resistant to the stress. Investigations such as these indicate that it is possible by proper application of experimental methods not only to study those conditions which are essential to the production of behaviour disorders, but also to isolate those attributes of the environment and of the organism which play contributory parts.

disorganizations of behaviour. Such studies as these have pointed to certain conditions which are critical to the production of behaviour disorders but not to all the essential conditions. For instance, it is possible for the conditions I have been discussing to be present and yet for no disorganization of behaviour to appear; the organism may 'leave the field' of the conflict or frustration. Rather early in the research in experimental psychopathology it was demonstrated that the organism's possible responses must be restricted if breakdown of normal behaviour is to occur. Two investigators (Anderson and Liddell, 1935), using animals as subjects and employing the Pavlovian differential conditioning technique for producing behaviour disorders, reported that subjects confined by a harness became restless, showed disturbances in respiration and other somatic functions, and developed aberrant behaviour patterns both inside and outside the laboratory, while subjects not so restricted revealed no such symptoms. Another study (Karn, 1940) concluded that patterns of learned behaviour may impose limitations on the repertoire of responses available for compensatory activities and thus contribute to conditions obstructing adjustment. 'Alternative responses may not be available because through habit the organism's activity is restricted to a definite response sequence.'

Other attributes of the environment and of the organism may contribute to the development of behaviour disorders without actually being essential conditions. A number of such contributing causes are associated with the concept of 'stress tolerance' or 'stress threshold'. Many experimental studies have demonstrated that the behaviour of some individuals is rapidly disorganized under stress, while other individuals are very resistant to breakdown. It is reasonable to suggest that behaviour disorders can be produced in all living organisms provided the stress to which they are exposed is of sufficient intensity and duration but this hypothesis is still in need of verification. A recent experiment in our laboratory (Ainsworth, 1953) lends partial support to its validity in that the number of human subjects showing definite signs of disorganization under stress increased as the intensity of the stress increased. So striking are

as possible from the immediate source of the stress. In other instances the subject may revert to a type of response pattern which characterized its behaviour at a much earlier period in its life history—a form of 'regression'. Repeated head-tossing, tic-like movements, fixed posturing and other apparently non-adaptive acts may develop. In general, attempts to measure behaviour when it takes these forms have been limited to two procedures. The simplest procedure involves the classification of the various behaviour patterns into different categories, each with its distinctive characteristics. The number of such categories and the frequency with which each appears under particular experimental conditions can be determined. Changes in experimental conditions may then be reflected in changes in the frequencies of the various categories. The second procedure is possible as soon as individual differences in behaviour are observed and subjects can be placed in a rank-order relative to the degree to which they evidence the behaviour or to the length of time over which the behaviour persists. When a behaviour pattern can be ranked in this way it is possible to make precise statements about its central tendency and the range of its variability, to determine the degree to which it is related to other variables, and to state the level of confidence with which the behaviour pattern differs in its expression between different groups of subjects. The application of this second procedure in experimental psychopathology increases the variety of different statements which can be made about behaviour disorders as well as the precision with which these statements can be made.

Instances in which the experimental psychopathologist studies the effects of exposure to stress by observing the ways in which later behaviour is affected are clearly illustrated in procedures involving the use of insoluble problems as a means of 'frustrating' human and animal subjects. After the subjects have been forced to respond in the insoluble problem situation a soluble problem is introduced. The number of trials taken by the experimentally frustrated subjects to reach a solution as compared with the performance of suitable control subjects becomes a measure of the effects of frustration.

From the very early observations of experimentally produced

The examples I have just given were selected to illustrate the extensive variety of independent variables with which the experimental psychopathologist is concerned, variables which he must be able to control and alter systematically. It is apparent that the experimenter exerts his control by manipulating various attributes of the physical and social environment in which the organism is placed or of the organism itself.

Dependent Variables

In experimental psychology the effects of manipulating independent variables are revealed, if at all, in the behaviour of the organism under observation. Hypotheses state the supposed nature of these effects and the effects themselves constitute the dependent variables which the investigator must systematically observe and, in some way, measure. The examples already given indicate the wide range of behaviour patterns exhibited by organisms exposed experimentally to various conditions of stress. Taken as a whole these behaviour patterns are similar in many ways to the 'symptoms' with which the clinical psychopathologist is concerned; in fact so similar that Pavlov referred to the behaviour of his dogs after breakdown in the conditioning situation as 'experimental neurosis'. Often the experimental psychopathologist describes his dependent variables directly in terms of the changes in behaviour produced by exposure to stress. In other instances he measures the effects exposure to stress may have on subsequent behaviour. Still other measures are obtained by observing changes in neurophysiological processes during and after the exposure.

A number of response characteristics have been used as signs that disorganization of behaviour was occurring during failure [illegible] These signs [illegible] sponding, appearance of rigidly stereotyped behaviour patterns, and attempts to leave the experimental field. Changes in level of general irritability and spontaneous activity may occur. 'Aggressive' actions towards other organisms and towards the experimental situation are often observed Frequently 'phobic' responses appear in which the subject struggles to retreat as far

in a series of studies (Hoagland, 1947; Hoagland, 1948) centred on the role of adrenal cortical hormones in adjustment to stress, some of the most exciting of the recent studies in experimental psychopathology. The investigators followed the experimental procedure of determining the output of adrenal cortical hormones by their subjects before, during, and after exposures to such stresses as flying airplanes, operating machines, exposure to heat and cold, taking examinations, and even under conditions no more severe than failing to reach their levels of aspiration in performing a simple motor task (Berkeley, 1952). All these conditions resulted in an increased output of the hormones in 'normal' subjects. However, mental hospital patients, when exposed to similar situations, showed little or no increase in hormone production. The researchers concluded (Hoagland, 1948): 'What is defective in the patient is his inability to increase his hormone output in response to environmental needs.' We are still in no position to state whether or not these associations between neurophysiological reactions and behaviour changes under stress are of cause and effect, but the research illustrates the way in which our knowledge of psychopathological states may be greatly enhanced by the study of neurophysiological as well as behavioural variables.

Research Design

The testing of hypotheses in experimental psychopathology requires the selection of appropriate scientific strategies or research designs which can then be put into effect by means of appropriate research techniques. Such strategies are based upon logical methods of reasoning which provide maximum confidence in the interpretation of research observations. Examination of the examples I have used to illustrate various key points in the methods employed by experimental psychopathologists shows that two principal strategies have seen major service. One requires comparisons of two sets of circumstances, usually referred to as control and experimental, in one of which the independent variable is present and, in the other, absent. If the results of the experiment indicate that the dependent variables under observation were present or absent in a corresponding

psychopathological states attention has been directed to the generalizations of behaviour disorders beyond the precise conditions under which they originally developed. Pavlov noted that the breakdown of behaviour in his dogs as a result of differential conditioning procedures soon spread to their behaviour outside the laboratory. Investigators using human subjects have observed that certain reactions to stress, for instance mild aggression, may continue to be apparent for some time after the experiment is over and may be elicited by stimuli with little if any relation to the experimental situation. Recently experimental analyses of the conditions affecting the generalization of behaviour disorders have been undertaken. The general approach in such analyses may be illustrated by reference to one experiment (Miller and Kraeling, 1952) in which rats, exposed to approach-avoidance conflict in a simple straightaway, developed characteristic vacillations of behaviour. Generalization of these effects of conflict was studied by testing some of the animals in the same straightaway in which they were trained, others in a straightaway differing in a few specific regards from the original, and still others in a very different straightaway. Animals tested in all three situations showed vacillation, indicating generalization of the behaviour disorders. By varying systematically different relevant attributes of the organism and of the environment, experimental answers can be given to questions regarding the conditions upon which the nature as well as the aetiology of behaviour disorders are based.

Important contributions to experimental psychopathology have been made by investigators who have focused their attention on the widespread changes in the absolute level and in the variability of neurophysiological activities which are correlated with the appearance of behaviour disorders under stress. One [illegible] title [illegible] has [illegible] nges generally follow—of a 'general-adaptation-syndrome'. These changes reflect specific and non-specific actions of the autonomic nervous system and the endocrine system. The intimate association of such actions with changes in behaviour is reflected

have been operating during his observations. The tests make it possible for him to state a level of confidence which he can place in any differences or concomitant variations he may have observed. He sets himself very rigorous standards for this level of confidence and usually accepts a convention which permits him to conclude that his results are significant only if they could be expected to occur by chance in 5 per cent or less of his observations.

There is no scientific strategy which does not have its limitations as well as its advantages when applied to problems of psychopathology. In considering these limitations it becomes apparent why *experimental* psychopathology should not claim to be the *single* desirable approach to the study of behaviour [illegible] observational methods'. For very obvious reasons the experimental production of extreme psychopathological states in man is not feasible. In practice this has meant that experimental studies involving man as a subject have been limited to the temporary disorganization of behaviour and not the development of chronic behaviour disorders. It may be argued that the [illegible] aetiology and nature of both. Temporary behaviour disorganization and extreme psychopathological states may represent ends of a continuum and not some clearly definable dichotomy.

This limitation on the use of man as a subject for research has led many investigators to take a comparative phyletic approach in their work. Such an approach is not unusual in the biological sciences contributing to medical research. It is based on the assumption that man is part of an evolutionary schema and not an organism entirely distinct in biological and psychological characteristics from his infrahuman relatives The experimental psychologist who employs a comparative phyletic

manner, then it can be said with confidence that the independent and dependent variables are related. In several of the experiments previously described one group of subjects was exposed to certain conditions under which disorganization of behaviour occurred, while control subjects not so treated developed no unusual symptoms. It was then possible to argue that the conditions in question were associated with the production of the behaviour disorders observed. The second strategy illustrated by my examples requires that the independent variable be altered in degree or magnitude. If, when an independent variable is altered systematically, a certain dependent variable is observed to change concomitantly, it can be said with confidence that the two variables are related and the degree of this relation can be calculated. Use of this strategy is illustrated by the experiment in which increases in the intensity of the stress to which different groups of subjects were exposed were associated with increases in the numbers of subjects showing signs of behaviour disorganization.

Both these strategies require that all relevant conditions other than those constituting the independent variable be controlled. Otherwise no conclusions regarding relations between independent and dependent variables can be drawn. In psychology, where so many conditions may act upon the dependent variables under observation, the exercise of this 'principle of control' is often a very difficult task. Individual differences in organisms, inequalities of conditions to which they are being or have been exposed, and changes during observation due to motivation, practice, fatigue, and so on, must all be taken into consideration. Techniques have been developed for eliminating potentially confusing elements, for holding them constant while varying the experimental variables, for randomizing their occurrence, and for measuring them in such a way that they may be removed by appropriate statistical techniques in analysing the results of observations.

In analysing his results the experimental psychologist begins by assuming that any apparent relations between variables may have occurred by chance. He subjects his data to statistical tests which allow him to estimate the extent to which chance may

cerned with only one approach to the study of problems in psychopathology, an approach which takes as its model the rigorous requirements of experimental science. I hope that the picture I have painted has given a general understanding of the requirements and limitations of this approach. When it can be applied the experimental approach is likely to provide the most accurate and precise answers to the problems of psychopathology.

REFERENCES

AINSWORTH, L. (1953). Ph.D Thesis, University of London
ANDERSON, O. D and LIDDELL, H S. (1935). *Arch Neurol Psychiat* 34, 330
BARKER, R. G. (1942). 'An Experimental Study of the Resolution of Conflict in Children' In McNemar, Q and Merrill, M. A., *Studies in Personality*. McGraw-Hill, New York, N.Y.
BARKER, R. G., DEMBO, T. and LEWIN, K (1941) *University of Iowa Studies in Child Welfare*, 18, 1.
BERKELEY, A. W (1952). *J. comp physiol. Psychol* 45, 443.
FREUD, S. (1920). *A General Introduction to Psychoanalysis* Liveright Publ. Co., New York, N Y
GANTT, W. H. (1944). *Psychosomat. Med. Monogr* 3, 1
HALL, C. S. (1938) *Sigma Xi Quart.* 26, 17
HOAGLAND, H (1947) *J. comp. physiol. Psychol.* 40, 107.
HOAGLAND, H (1948) *Research Reviews*, Office of Naval Research, Washington, D C.
HUNT, J McV. (1941) *J abnorm soc. Psychol* 36, 338.
KAHN, H. (1940) *J. gen Psychol* 22, 431
LEVY, D M (1934) *Amer J Ortho-psychiat* 4, 203.
MAIER, N R F (1949) *Frustration The Study of Behaviour without a Goal* McGraw-Hill, New York, N.Y.
MASSERMAN, J H (1943) *Behaviour and Neurosis An Experimental Psychoanalytic Approach to Psychobiologic Principles.* University of Chicago Press, Chicago, Ill.
MILLER, N (1944). 'Experimental Studies of Conflict.' In Hunt, J. McV, *Personality and the Behavior Disorders*, Ronald Press, New York, N Y.
MILLER, N. and KRAELING, D. (1952) *J exp. Psychol* 43, 217
PAVLOV, I P (1928) *Lectures on Conditioned Reflexes.* International Publ Co., New York, N Y
PAVLOV, I P (1941) *Lectures on Conditioned Reflexes: Conditioned Reflexes and Psychiatry* International Publ Co, New York, N Y

approach designs his research to '... involve the *extensive* observation of homologous behaviour patterns in a *variety* of forms or types of organisms' (Russell, 1951). This approach may lead to the discovery of general principles which cannot be seen clearly in observations of only a single species. To those experimental psychologists whose interests are primarily anthropocentric infrahuman animals may be used in research as substitutes for man—as research tools. It is very true that the generalization of principles discovered during research on infrahuman animals to man involves some precarious logical manoeuvres, but the examples where such generalization has proved warranted are numerous. Pavlov expressed an extreme attitude towards the role of infrahuman animals in the experimental study of psychopathological states when he wrote (Pavlov, 1941): 'I am convinced that the decision, or the conditions favourable to a *decision, of many important questions of etiology, the natural* systematization, the mechanism and finally the treatment of neuroses in the human being lie in the hands of the animal experimenter.' Our present knowledge in experimental psychopathology, as in experimental psychology generally, derives much from research on infrahuman animals.

Concluding Remarks

With estimates from several countries indicating that the number of hospital beds occupied by patients suffering from behaviour disorders may be greater than the number occupied by patients with all other diseases combined, there can be little doubt of the important and urgent need for research into the aetiology, *nature, and treatment of psychopathological states.* The field is a large one and provides adequate scope for anyone who cares to explore it regardless of the specific discipline to which he subscribes or the particular approach he favours. There is a place for those who have the perseverance to undertake the multitude of detailed, time-consuming investigations that eventually form the basis for general principles or laws, as well as for those who are capable of contributing new ideas and developing new techniques. If I have succeeded in making myself at all clear, it must be apparent that I have been con-

COMPLETE LIST OF LECTURES

* *Lectures included in the published volumes*

1951–52

*The Scientific Approach to Medical Research
PROFESSOR E. ADRIAN, O.M., M D , F.R.C.P., F.R.S. (*Cambridge*)

The Philosophy of Science
PROFESSOR H. DINGLE, D.SC., A.R.C.S. (*London, University College*)

Studies by X-ray Analysis, Electron Microscopy and Supporting Techniques of the Structure of Biological Macromolecules and the Tissues formed from them
PROFESSOR W. T. ASTBURY, M.A., SC D , F.R.S. (*Leeds*)

*Tissue Responses to Injury
PROFESSOR G. R CAMERON, M B , B.S , D.SC , F.R C.P , F.R S. (*London, University College Hospital Medical School*)

*Studies of Normal and Pathological Physiology of the Kidney
D VAN SLYKE, M.D (*New York, U.S A*)

Biochemical Function of Vitamin B_1
PROFESSOR R A. PETERS, M.D., F R S. (*Oxford*)

Nutritional Science and the Feeding of Populations
PROFESSOR S J COWELL, M B , F.R.C P.
(*London, University College Hospital Medical School*)

*Nutritional Assessment of the Individual
PROFESSOR J. YUDKIN, M A , M.D , PH.D.
(*London, Queen Elizabeth College*)

Folic Acid and Vitamin B_{12} in the Metabolism of Micro-organisms
D D WOODS, M.A., PH.D. (*Oxford*)

RUSSELL, R. W. (1951). *The Comparative Study of Behaviour*. H. K. Lewis, London.
RUSSELL, R. W. (1953). *Internat. J. Psychoanal.* **34**, suppl, 1.
SELYE, H. (1950). *The Physiology and Pathology of Exposure to Stress*. Acta, Montreal, Canada.

*The Physiological Basis of Visual Sensation
L. C. Thomson, Ph.D. (*London, Institute of Ophthalmology*)

*The Contribution of Science to the Practice of Health in the First Quarter of the Twentieth Century
Professor J. M. Mackintosh, M.D., F.R.C.P., LL.D.
(*London School of Hygiene and Tropical Medicine*)

Methods of Research in Social Medicine
Alice M. Stewart, M.A., M.D., F.R.C.P. (*Oxford*)

Environmental Factors in Work
Professor G. P. Crowden, O.B.E., T.D., D.SC.
(*London School of Hygiene and Tropical Medicine*)

Methods of Research in Industrial Medicine
Donald Hunter, M.D., F.R.C.P. (*London Hospital Medical College*)

*Recent Light on Mammary Function
Professor H. D. Kay, C.B.E., D.SC., F.R.S. (*Reading*)

*Research on Ageing
Professor E. C. Dodds, M.D., F.R.C.P., F.R.S.
(*London, Middlesex Hospital Medical School*)

Carcinogenesis
Professor A. Haddow, M.D., D.SC., PH.D.
(*London, Institute of Cancer Research*)

Physical Techniques in the Medical Applications of Ionizing Radiations
Professor W. V. Mayneord, D.SC.
(*London, Institute of Cancer Research*)

*Biological Effects of Radiation
J. F. Loutit, D.M., M.R.C.P. (*M.R.C., Harwell*)

*The Effects of Haemorrhage on the Cardiovascular System in Man
O. G. Edholm, B.SC., M.B., B.S.
(*National Institute for Medical Research, Mill Hill*)

*Lability of Blood Pressure
Professor M. L. Rosenheim, M.A., M.D., F.R.C.P.
(*London, University College Hospital Medical School*)

Cardiac Output in Man (*film*)
Professor J. M. McMichael, M.D., F.R.C.P.
(*Postgraduate Medical School of London*)

Aspects of Bacteriostasis
PROFESSOR B. C. J. G. KNIGHT, D.SC. (*Reading*)

*Some Aspects of Antibacterial Immunity
A. A. MILES, M.A., M.B., F.R.C.P.
(*National Institute for Medical Research, Mill Hill*)

*Antibodies and Immunity to Virus Infection
PROFESSOR A. W. DOWNIE, M.D., D.SC. (*Liverpool*)

Current Knowledge on Nature of Viruses
F. C. BAWDEN, M.A., F.R.S. (*Rothamsted Experimental Station*)

*Viruses as the Causes of Diseases
PROFESSOR S. P. BEDSON, M.D., D.SC., F.R.C.P., F.R.S.
(*London Hospital Medical College*)

*Adrenal Hormones and ACTH
PROFESSOR F. G. YOUNG, M.A., D.SC., F.R.I.C., F.R.S. (*Cambridge*)

Chemical Aspect of Antibiotics
SIR ROBERT ROBINSON, O.M., D.SC., LL.D., F.R.S. (*Oxford*)

*Growth of the Human at the Time of Adolescence
J. M. TANNER, M.D. (*London, St. Thomas's Hospital Medical School*)

Bone Growth and Repair
PROFESSOR W. D. NEWCOMB, M.D., F.R.C.P.
(*London, St. Mary's Hospital Medical School*)

The Rate and Materials of Blood Formation
PROFESSOR L. J. WITTS, D.M., F.R.C.P. (*Oxford*)

Alimentary Factors in Blood Formation
PROFESSOR L. J. WITTS, D.M., F.R.C.P. (*Oxford*)

*Blood Coagulation in Theory and Practice
R. G. MACFARLANE, M.D. (*Oxford*)

The Psychology of Personality
PROFESSOR SIR CYRIL BURT, M.A., D.SC. (*London, University College*)

*On the Organization of Cortical Mechanisms
PROFESSOR SIR GEOFFREY JEFFERSON, M.S., F.R.C.P., F.R.C.S., F.R.S.
(*Manchester*)

Involuntary Nervous System (*2 lectures*)
PROFESSOR G. L. BROWN, C.B.E., M.SC., F.R.S.
(*London, University College*)

*The Life History of the Malaria Parasite
PROFESSOR P. C. C. GARNHAM, M.D., D.SC., D.P.H.
(*London School of Hygiene and Tropical Medicine*)

*Human Haemoglobins
J. C WHITE, M.B., CH.B. (*Postgraduate Medical School of London*)

Uses and Applications of Human Blood Plasma Fractions
SIR ALAN DRURY, C.B.E., M.D., F.R S. (*Cambridge*)

*Preservation of Living Cells at Low Temperatures
A. S. PARKES, M.A., SC D., F.R S.
(*National Institute for Medical Research, Mill Hill*)

*The Life-span of Red Blood-cells
P. L. MOLLISON, B.A., M D. (*Postgraduate Medical School of London*)

*Biochemical Genetics
A. NEUBERGER, PH D., M.D , F.R S.
(*National Institute for Medical Research, Mill Hill*)

*Chromatography in the Study of Amino-Acid Metabolism
C. E DENT, B SC , PH.D., M D., F.R.I.C., M R.C P.
(*London, University College Hospital Medical School*)

*Selective Toxicity (*2 lectures*)
PROFESSOR ADRIAN ALBERT, D SC , PH.D., F R.I.C.
(*Australian National University*)

*Recent Progress in Antibiotics
PROFESSOR SIR ALEXANDER FLEMING, M.B., B.S , F.R.C.S., F R C P , F.R.S. (*London, St. Mary's Hospital Medical School*)

*Virus Adaptability in Relation to Human Disease
PROFESSOR WILSON SMITH, M D., F.R S
(*London, University College Hospital Medical School*)

*Carcinogenesis
P. R. PEACOCK, M B., B.S , F R.F P S G. (*Glasgow*)

*The Metabolism of the Adrenocortical Hormones
PROFESSOR G F MARRIAN, D SC , F R.S (*Edinburgh*)

The Pituitary Gland and Reproduction
S J. FOLLEY, D SC , PH D., F.R S. (*Reading*)

*The Physiology of Parturition
PROFESSOR A. ST G HUGGETT, D SC , PH.D.
(*London, St. Mary's Hospital Medical School*)

1952–53

The Influence of Language on Medicine
PROFESSOR J. Z. YOUNG, M.A., F.R.S. (*London, University College*)

The Nature and Place of Thinking in Medicine
PROFESSOR SIR FREDERIC BARTLETT, C.B.E., M.A., F.R.S. (*Cambridge*)

*The Methodology of Clinical Science
PROFESSOR SIR JAMES SPENCE, M.C., M.D., F.R.C.P. (*Newcastle*)

The Statistical Approach
PROFESSOR BRADFORD HILL, C.B.E., D.SC., PH.D. (*London School of Hygiene and Tropical Medicine*)

Control of Respiration
PROFESSOR C. G. DOUGLAS, C.M.G., M.C., M.A., M.D., F.R.S. (*Oxford*)

Life at High Altitudes
PROFESSOR SIR B. H. C. MATTHEWS, C.B.E., M.A., SC.D., F.R.S. (*Cambridge*)

*The Physiological Effects of Gravity
GROUP CAPTAIN W. K. STEWART, C.B.E., A.F.C., R.A.F. (*Farnborough*)

The Kinetics of Rapid Chemical and Biological Reactions
PROFESSOR F. J. W. ROUGHTON, M.A., PH.D., F.R.S. (*Cambridge*)

Metal Catalysis and Intracellular Respiration
PROFESSOR D. KEILIN, F.R.S. (*Cambridge*)

Biophysical Studies of Connective Tissue
PROFESSOR J. T. RANDALL, D.SC., F.R.S. (*London, King's College*)

*The Functional Significance of Connective Tissue
A. H. T. ROBB-SMITH, M.A., M.D., F.R.C.P. (*Oxford*)

*Silicosis
PROFESSOR E. J. KING, PH.D., D.SC. (*Postgraduate Medical School of London*)

Pulmonary Asbestosis
PROFESSOR MATTHEW STEWART, C.B.E., F.R.C.P. (*Leeds*)

Organ Culture in Biological and Medical Research
HONOR B. FELL, F.R.S. (*Cambridge*)

The Biology of the Foetal Membranes and the Placenta
PROFESSOR E. C. AMOROSO, M.B., B.CH., PH.D.
(London, Royal Veterinary College and Hospital)

*The Physiological Actions of the Sex Hormones
P. M. F. BISHOP, B CH., D.M.
(London, Institute of Obstetrics and Gynaecology)

*Acid and Alkaline Phosphatase in Disease
PROFESSOR E. J. KING, M.A., D.SC., PH D., F.R.I.C.
(Postgraduate Medical School of London)

Principles of Chromatography
R. L. M. SYNGE, F.R.S. *(Aberdeen)*

Some Recent Ideas about the Proteins and other Biological Macromolecules
PROFESSOR W. T. ASTBURY, M.A., SC.D., F.R.S. *(Leeds)*

*Stress and Thyroid Activity
PROFESSOR G. W HARRIS, SC.D., M.D., F.R.S.
(London, Institute of Psychiatry)

The Vascular Lesions of Pituitary Necrosis
PROFESSOR H L SHEEHAN, D.SC., M D, F.R.C.P, F.R.C.O.G.
(Liverpool)

*Science and History
H. E. SIGERIST, M D., LL B., D.LITT. *(Yale, U.S.A.)*

*Anticholinesterases
W. D. M. PATON, M A, B M, B.CH.
(London, University College Hospital Medical School)

*The Scientific Approach to Dermatology
R M. B. MACKENNA, M.A., M.D., F.R.C.P.
(London, Institute of Dermatology)

*Hemispherectomy and the Localization of Function
E. A. CARMICHAEL, C.B.E, F.R C.P.
(London, Institute of Neurology)

*Tissue Repair
PROFESSOR G. R. CAMERON, D SC., F.R.C.P., F.R.S.
(London, University College Hospital Medical School)

The Design of Clinical Experiments
D. D REID, M D., PH.D.
(London School of Hygiene and Tropical Medicine)

*The Physiology of the Autonomic Nervous System
W. S. FELDBERG, M.D., F.R.S.
(*National Institute for Medical Research, Mill Hill*)

*The Principles of Gangilonic Block
W. D. M. PATON, M.A., B.M., B.CH.
(*London, University College Hospital Medical School*)

*Cholinesterases and Anti-Cholinesterases
PROFESSOR R. H. S. THOMPSON, D.M.
(*London, Guy's Hospital Medical School*)

Certain Aspects of Nerve Repair
H. J. SEDDON, M.A., D.M., F.R.C.S. (*London, Institute of Orthopaedics*)

Development of the Heart in Relation to Congenital Heart Disease
PROFESSOR J. D. BOYD, M.A., M.SC., M.D. (*Cambridge*)

Dynamics of the Heart Beat
PROFESSOR A. HEMINGWAY, M.SC., M.B., CH.B. (*Leeds*)

The Natural History of Essential Hypertension
PROFESSOR G. W. PICKERING, M.A., M B., F R.C P.
(*London, St. Mary's Hospital Medical School*)

Causes of Hypotension in Man
PROFESSOR E. P. SHARPEY-SCHAFER, M.B., B.CHIR., F.R.C.P.
(*London, St. Thomas's Hospital Medical School*)

1953–54

Scientific Method in Medical Research
SIR HENRY DALE, O.M., G.B.E., F.R.C.P., F.R.S.

*Experimental Psychopathology
PROFESSOR R. W. RUSSELL, M.A., PH.D., D.SC.
(*London, University College*)

*Biological Synthesis
G. POPJÁK, M.D.
(*National Institute for Medical Research, Mill Hill*)

*Acetylcholine and the Maintenance of the Cardiac Rhythm
PROFESSOR J. H. BURN, M.A., F.R.S. (*Oxford*)

*The Growth Hormone of the Anterior Pituitary Gland
PROFESSOR F. G. YOUNG, M.A., D.SC., F.R.I.C., F.R.S. (*Cambridge*)

*The Genetics of Some Biochemical Abnormalities
PROFESSOR J. B. S. HALDANE, M.A., D.SC., F.R.S.
(London, University College)

*The Chemotherapy of Cancer
PROFESSOR E. BOYLAND, PH.D., D.SC.
(London, Institute of Cancer Research)

*Causes of Failure in Antibiotic Therapy
PROFESSOR L. P. GARROD, M.D., F.R.C.P.
(London, St. Bartholomew's Hospital Medical College)

*Reactions to Bacterial Invasion
PROFESSOR A. A. MILES, C.B.E., M.D., F.R.C.P.
(London, Lister Institute of Preventive Medicine)

*The Action of Bacterial Enzymes on Immunizing Antigens
PROFESSOR LORD STAMP, M.A., M.B., CH.B.
(Postgraduate Medical School of London)

*Antiviral Immunity
PROFESSOR H. B. MAITLAND, M.SC., M.D. *(Manchester)*

*Antimalarial Drugs
C. C. CHESTERMAN, O.B.E., M.D., F.R.C.P.
(London, Middlesex Hospital Medical School)

Chemotherapy of Tuberculosis
P. M. D'ARCY HART, M.A., M.D., F.R.C.P.

Blood Groups
R. R. RACE, PH.D., M.R.C.S., F.R.S.
(London, Lister Institute of Preventive Medicine)

Congenital Haemolytic Anaemia
J. V. DACIE, M.D., M.R.C.P.
(Postgraduate Medical School of London)

*The Supporting System and its Disorders
PROFESSOR J. H. KELLGREN, M.B., B.S., F.R.C.S., F.R.C.P.
(Manchester)

*Regulation of Body Water
PROFESSOR G. M. BULL, M.D., M.R.C.P. *(Belfast)*

Plate I



Fig 1 Deposition of bone salts in epiphyseal cartilage, by action of phosphatase on phosphoric ester. *Upper left* Tibia of a rachitic rat *Lower left* Same tibia (*serial slice*) immersed in solution of phosphoric ester (calcium glycerophosphate) Note extensive calcification *Upper right* Humerus of rachitic rat immersed in solution of inorganic calcium phosphate. Note slight calcification *Lower right* Tibia immersed in solution of phosphoric ester (calcium glycerophosphate) Note extensive calcification (after Robison, 1932)

(*See p* 191)

Plate II



	Femur	Tibia	Humerus
Shaft	6·3	1 4	1 4
Metaphysis	4 5	1 6	3 9
Epiphysis	0 4	0 2	0·1

(See p 192)

Plate III

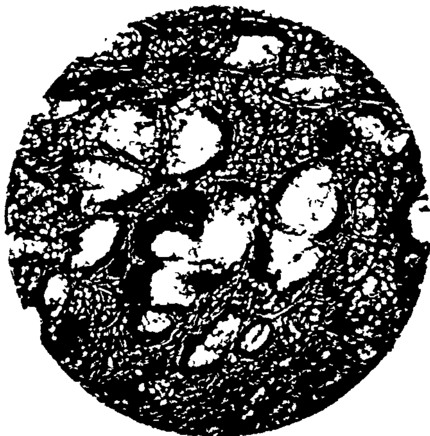
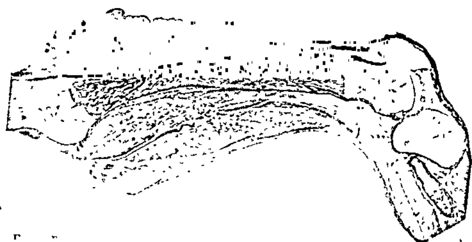


Fig 6. Histological demonstration of acid phosphatase in the prostate gland. Stained by Gomori's technique (see Pearse, 1953).

(See p. 204)

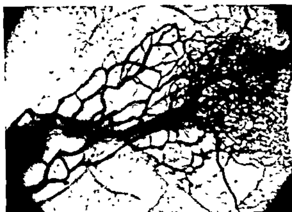
PLATE II



	Femur	Tibia	Humerus
Shaft	6 3	1 4	1 4
Metaphysis	4 5	1 6	3 9
Epiphysis	0 4	0 2	0 1

(See p 192)

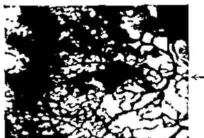
PLATE V



a



b



c

under the low injection pressure, has filled with ink.

(*c*) A 24-hour-old micro-abscess due to *Strep. pyogenes*. The abcess is the grey opacity in the centre of the photograph; and the plexus is interrupted by blocked lymphatic vessels over its centre (Approx × 10)

(*See p* 249–50)

Plate IV



Fig. 9. Prostate gland stained for acid phosphatase (Gomori's method), and showing cystic cavities which were filled with prostatic plasma of high acid phosphatase activity (6,000 units/ml)

(See p 211)



Fig 1 Statue of St. Bartholomew in Milan Cathedral (By courtesy of Mr Ogier Ward.)

(See p 350)



FIG 6 Acute eczema affecting the limbs, illustrating 'sympathy' and 'homology'

(See p 366)

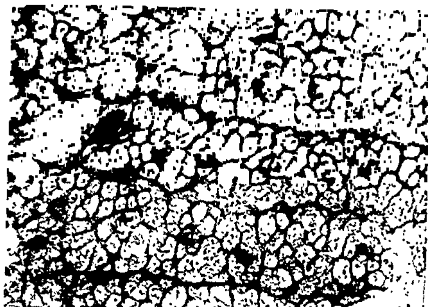


FIG 2. Skin of human leg, × 24. Stained with anthracene blue. 500μ horizontal section To show rete Malpighii seen from below (reproduced from Leach, 1952).

(See p 351)



FIG 4. Infected seborrhoeic dermatitis imitating impetigo contagiosa.

(See p 357)

PLATE IX



FIG. 7 Lichen planus, illustrating 'sympathy'

(See p 366)